Facilitating Childbearing in Taiwan: The Role of Domestic Gender Equity and Parental Leave

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor in Philosophy by

Jac Arthur Celyn Thomas

Department of Geography and Planning
University of Liverpool
United Kingdom

August 2021

Abstract

Across many high- and middle-income countries, fertility rates have declined to unsustainably low levels. This phenomenon is particularly acute in Taiwan, where fertility has been below the "lowest-low" level of 1.3 births per woman since 2003 and hit the extremely low level of 0.895 in 2010. These trends may lead to long-term social and economic problems due to the narrowing of the working-age population and increases in old-age dependency ratios, and women are having fewer children than they desire. Ultimately, the persistence of below-replacement fertility will lead to population decline.

On the academic side, theorists have attempted to explain low fertility in Taiwan in terms of gender inequity, arguing that gender-equitable career ambitions and gender-inequitable family expectations have placed a double burden on women who want to have children. On the policy side, the Taiwanese government has attempted to increase fertility through pro-natal policies, the most expensive of which has been the Parental Leave Allowance (PLA) introduced in 2009. There is heretofore no micro-level statistical evidence that traditional familial and domestic obligations reduce women's fertility in Taiwan. There is also no research attempting to evaluate the causal impact of the PLA or any other national pro-natal policy on fertility in Taiwan. Moreover, the wider academic literature is unclear on whether parental leave impacts fertility, and is methodologically unclear on how to identify the effect of leave policies on fertility.

In this thesis, we aim to find out whether gender inequity is a cause of low fertility in Taiwan, and whether parental leave policies have an impact on fertility. This thesis is structured around three research papers: an empirical evaluation of the effect of housework division on realised fertility; a systematic review of the effects of parental leave policies on fertility; and an empirical evaluation of the causal impact of the PLA on fertility. In the first and third papers, we use econometric methods to predict next births based on housework division and PLA eligibility respectively. In the second paper, we use systematic review methods to find and synthesise all the best available evidence of the causal impact of leave on fertility.

In the first paper, we find that the division of housework between married couples of parity 1 or higher has a large impact on subsequent fertility, with more equal divisions being associated with higher fertility. In the second paper, we develop a new conceptual and methodological framework to decompose the different impacts of leave policies on fertility. We also argue that certain types of effects are more informative to pro-natal policy-makers than others. Surveying the available evidence in terms of our novel framework, we find that studies

identifying a broad class of effects consistently report positive relationships, and those reporting null relationships are only identifying a narrow class of effects of marginal interest to policy-makers. However, the findings are restricted to the Western countries in which these studies have been conducted. In the third paper we use our framework to evaluate the causal impact of the PLA on fertility. Contrary to our findings in Western settings, we find that the PLA did not increase fertility for women who were always working, but may have increased second births among women who had been in and out of work since marriage.

Our findings make several contributions to research on low fertility, both globally and in Taiwan. Firstly, we empirically confirm the validity of gender equity theory as an explanation of low fertility in Taiwan. Secondly, we provide an intelligible framework for analysing the effects of pro-natal policies on fertility, that reflects the imperatives of policy-makers. Thirdly, we show that parental leave policies have a consistently positive impact on fertility in Western societies, and fourthly we show that the PLA had a limited impact on fertility in Taiwan, despite its cost.

Acknowledgements

I would like to thank my supervisors, Dr Francisco Rowe and Professor Eric S. Lin. Prior to my enrolment I believed myself to have a decent understanding of the processes and standards of social research, and Dr Rowe and Professor Lin quickly disabused me of that belief. I would not have been able to complete my research to its current standard without their instruction. In Liverpool, Dr Rowe's feedback on every paper and chapter has improved my work beyond recognition, and ensured it meets the standards for publication. In Hsinchu, Professor Lin's expertise in econometrics enabled me to develop the conceptual and methodological thinking that runs through the three papers. Both supervisors also advocated on my behalf in their respective institutions, helping me tangle with the rare and perplexing bureaucratic challenges associated with studying in two countries. On a personal level, both have looked out for my financial and emotional well-being through what at times has been somewhat of a cash-strapped and isolating experience. More positively, these past 4 years have really broadened my horizons and been enormously enriching in terms of my academic, professional, and personal lives. Francisco and Eric: you made this possible, so I thank you both as mentors and friends.

I would also like to thank my academic colleagues at NTHU and UoL, as well as external academics who have given me their time. Particular thanks goes to Dr Paul Williamson, who co-authored Paper 2 and provided incisive feedback on focus and argumentation. Thanks also goes to Professor Min-Teh Yu, who gave feedback on Paper 1 at the WEAI conference in San Francisco. I thank Professor Stuart Gietel-Basten for an illuminating discussion that sparked several key ideas about the methods for evaluating policies. At the UoL library, I would like to thank Ms Lisa Hawksworth for advice on the search strategy in Paper 2. At the institutional level, the Geographic Data Science Lab in Liverpool and the Department of Economics in NTHU have been welcoming and stimulating environments in which to exchange ideas and conduct my research, and I thank both of these communities. I would also like to thank the departmental and programme administrators in Liverpool and Hsinchu, who have been incredibly helpful in facilitating the dual-program PhD.

On a personal level, I would like to thank my friends and family for their kindness and support. I arrived in Taiwan with no personal connections, but quickly found a community in the Department of Economics. Particular thanks goes to Li-Dan Shang, my friend and colleague on the dual-PhD programme, who helped me adjust to life in Taiwan. Through teaching English part-time, I became acquainted with the Zhang, Kuo, and Huang families, all of whom

became close personal friends and provided a bedrock of emotional stability for what might have otherwise been a fairly rootless and temperamental three years in Taiwan. The Zhang family particularly – Eric, Carol, and 弟弟 Joe – became like a second family to me. They took me on holidays, cared for me in hospital when I had appendicitis and allergic reactions, and were incredibly generous. In the UK I would like to thank my parents Elaine and Kevin, and my brother Joseff, who have always encouraged my education and supported me in my personal life. They have been particularly supportive over lockdown when I have mostly been at home, bringing me food, proofreading, and tolerating the odd hours I sometimes keep. I would also like to thank them for progressing from parity 1 to parity 2, helping to sustain replacement-level fertility in the UK and enabling me to exist. It is to my family that I dedicate this thesis, as well as to my adoptive family in Taiwan.

Finally, I would like to thank my PhD examiners: Professor Stuart Gietel-Basten, Professor Yu-Hua Chen, Dr Frances Darlington-Pollock, Professor Stephen Jui-Hsien Chou, and Dr Jamie Goodwin-White. The *viva voce* exam was an intense experience at times, but I very much appreciated the degree to which the examiners had engaged with the assumptions and arguments of the thesis. I feel incredibly fortunate that the external examiners – Professors Gietel-Basten and Chen – agreed to examine the PhD, since they are some of the leading international researchers on Taiwanese and East Asian fertility. I really welcomed the robustness of their criticism, and sincerely want to thank them for the time, effort and energy they devoted to examining the work. Their feedback helped improve the framing and argument of the thesis immeasurably. Any remaining mistakes or idiosyncratic arguments – particularly regarding arguments around the "problem" of low fertility – are entirely my own.

As the sound of the playgrounds faded,
the despair set in.

Children of Men

95 per cent of economics
is common sense.

Ha-Joon Chang

Preface

Demography concerns itself with phenomena central to the human experience, such as birth, death, sex, ageing, gender, sexuality, family, relationships, and migration. My focus has always been on fertility, which at the highest level can be seen as part of an unbroken chain of reproduction stretching back to the first lifeforms on Earth. At the macro level, demography examines abstract collections of people – bounded politically, geographically, ethnically, or otherwise – called "populations." Populations persist through time even after all of its current members have died, through reproduction. The subjects and abstractions of demography can make the discipline very engaging to researchers with a philosophical inclination.

This PhD project began under the heading, "Why has fertility declined to a very low level in Taiwan?" Prior to the PhD, my masters thesis had focussed on an opposite question: why is fertility in sub-Saharan Africa (still) so high? Both questions are seemingly straightforward, but actually contain deep assumptions about the levels of aggregate fertility to be expected from different types of societies. Moreover, the political and economic contexts in which these questions are asked pose additional, normative assumptions about what levels of fertility are desirable. Usually, such normative assumptions are associated with restrictive moral prescriptions for sexual and reproductive behaviours, and particularly for women.

In most of social science, our aim is to describe or explain social phenomena free of any personal judgements or normative prescriptions. However, the subject matter of demography can make it very difficult – and even undesirable – to disentangle normative considerations from analysis. Historically and into the present, fears about demographic phenomena have provided the *animus* for innumerable violations of human rights and even atrocities, such as eugenics programmes, forced sterilisations, and ethnic cleansing. In some low fertility countries today, governments have begun to introduce authoritarian and traditionalist family policies aiming to reverse gender equality, and to encourage women to remain at home in order to have children. This is the moral universe in which demographers of low fertility conduct their research, and it is naive to discuss policies aiming to increase fertility without considering these normative factors.

A thorough exploration of the normative and foundational assumptions underlying analyses of fertility is beyond the scope of this thesis, and outside my expertise. Nonetheless, our analysis has to begin by making some defensible normative assumptions. This thesis begins with the assumptions that the continued existence of Taiwanese society is a desirable thing, and that governments should

only try to increase fertility by removing barriers to childbearing (rather than by coercing the behaviour of women). Along with the assumption that ageing and death will continue to deplete the population, I will argue that the first assumption implies that *at some point*, fertility will have to come back to near replacement level. Some commentators characterise discourses around low fertility as alarmist, and that the putative economic consequences of population ageing are dubious. But this begs the question: if fertility does not need to increase now, then when should it increase? Similarly, many commentators – including myself – criticise the inegalitarianism of some existing policies seeking to increase fertility. But again, this begs the question: what policies should we be using? Ultimately, the big question posed by Taiwan (and other low fertility countries) is this: how can post-transitional populations begin to manage their fertility levels in a way that will be sustainable in the long run?

The first epigraph of this thesis is an (admittedly emotive) evocation of the despair felt in a society that has ceased to reproduce itself. Clearly, Taiwan today is quite far from the apocalyptic situation depicted in *Children of Men*. However, that despair is not too different from the "emptiness" felt by those in the depopulating hinterlands of contemporary Eastern Europe (Dzenovska 2020). If low fertility rates persist, Taiwan may well experience a similar outcome.

Contents

List of Figures

List of Tables

1

Introduction

1.1 Low fertility

Across many high- and middle-income countries, men and women are not having enough children to replace themselves. More specifically, countries in Eastern, Southern, and Central Europe, and in East Asia, have experienced Total Fertility Rates (TFRs) below 1.7 for some decades now (Rindfuss et al. 2016; UN 2019). Fertility is particularly low in Taiwan, where in 2010, the TFR fell to 0.895 births per woman, the lowest fertility rate of any major country in history (UN 2019). Taiwan has had a TFR below the "lowest-low" level of 1.3 since 2003. Taiwan's fertility trend is comparable to other societies in East Asia, specifically South Korea, Japan, Singapore, and Hong Kong, societies which share a Confucian culture and comparable development trajectories (UN 2019). Taiwan's fertility trend can therefore be framed in a national or regional context (Lin and Kamo 2015).

1.2 Is low fertility a problem?

Broadly, there are three main clusters of arguments that low fertility either *causes* or *represents* negative social, economic, and demographic phenomena. The first two sets of arguments concern consequences of low fertility, and centre around population *ageing* and population *decline* respectively. The third set of arguments centre around people having fewer children than they would ideally like to have, and therefore conceptualises low fertility as representing a *failure of childbearing aspirations*. Some academics have been highly critical of much of this discourse, with notable critiques found in Gietel-Basten 2019, Coleman and Rowthorn 2011

and Kravdal 2010. Gietel-Basten 2019 takes a particularly firm stance against the idea that low fertility in East Asia even is a problem, arguing that we should conceptualise low fertility as simply the outcome of institutional malfunctions. In this section we explore these theories and debates, and make the argument that – *contra* Gietel-Basten – low fertility is a problem in and of itself (or at least lowest-low fertility is). We will make this argument on the basis of population decline and existential concerns.

In terms of population ageing, low fertility means that the absolute size of successive generations will decrease, reducing the size of the future labour force (Bloom et al. 2010; Poston and Zhang 2014; Sanchez-Romero 2013). The proportional decrease in the working-age population will increase the old-age dependency ratio, thereby increasing aggregate consumption relative to investment and obstructing long-term economic development (Bloom et al. 2011, 2010). In terms of population decline, long-term low fertility builds a negative momentum into the population structure due to the shrinking of childbearing-age cohorts, making long-term population decline likely (Feeney 2003). In the very long-term societies must find ways to enable people to sustain replacement-level fertility, or else face an existential threat.

At the individual level, low fertility often indicates that people are having fewer children then they would ideally like to have, representing a failure of childbearing aspirations (Basten and Verropoulou 2015; Brinton et al. 2018; Gietel-Basten 2019). Social surveys across low fertility countries in East Asia consistently find that most parity 0 women capable of childbearing aspire to have two children, and yet period and cohort fertility remain much lower than these aspirations (Casterline and Gietel-Basten 2018; Gietel-Basten 2019). Moreover, between 20-25% of women exit their childbearing years without having any children, numbers which are much higher than the percentage of women at the start of their childbearing years who say that want no children (Gietel-Basten 2018). In this way, the failure to realise childbearing aspirations can be seen as a problem for reproductive rights and life-fulfilment.

There are several counterarguments to the theories laid out above. As a counterpoint to fears about population ageing, the theory of the "Second Demographic Dividend" argues that an ageing population may increase savings rates, due to savings from (relatively large) cohorts born during the period of the demographic transition (Abío et al. 2017; Cai 2020; Mason and Lee 2006). In terms of the natural environment, Reher 2007 points out that population decline will likely have a positive impact on ecological systems. In a review of arguments on this topic, Kravdal 2010 argues that migration may be a lever for offsetting population declines in low fertility areas. Going even further, Gietel-Basten 2019 argues that

the whole conceptualisation of low fertility as a "problem" is mistaken. Rather, he argues that low fertility is the outcome of individuals being unable to meet their childbearing aspirations due to institutional factors, such as overwork, the cost of childrearing, and gender inequity. In this way, low fertility is simply a consequence of "upstream" problems in institutional arrangements. To support the argument, Gietel-Basten uses survey data to show that ideal family sizes are consistently higher than realised fertility across East Asia. Gietel-Basten also notes that the public conversation around low fertility has become toxic, with different generations and sexes all blaming each other. As an alternative, he argues for a rights-based approach to low fertility, which seeks to remove barriers to childbearing in order to allow individuals to meet their childbearing aspirations.

Despite the critiques of Gietel-Basten and others, we contend that lowest-low fertility is a problem in and of itself. To rehearse the arguments so far: concerns about population ageing, decline, and failures to realise childbearing aspirations are all used to conceptualise lowest-low fertility as a "problem." However, many commentators dispute the theories of the supposed negative socioeconomic consequences of lowest-low fertility, and even argue that we should regard lowest-low fertility merely as an outcome of other social problems. Here we make a minimal argument: that fertility in Taiwan has to increase at some point, otherwise the population will decline quite rapidly and eventually disappear. Assuming that death will continue to be a feature of human societies (and there is no reason to believe this isn't the case), it is simply a mathematical fact that if fertility rates in Taiwan are sustained at their present levels, the size of the population will decline dramatically in the long term. This is a very simple observation, and has no theoretical content – but this simplicity does not mean it isn't true. As a comparison, Gietel-Basten 2019 rightly highlights similarities between discourses of low fertility with popular (and often hysterical) fears of high fertility in the 1960s and 1970s. However, a very similar argument can be made about mid-century or historical transitional fertility declines: those fertility declines *had to have occurred at some stage*, assuming that mortality would remain low. Again, this argument has no theoretical content, but it is nonetheless a mathematical fact that populations cannot sustain large positive or negative growth rates in the long run. To paraphrase the Ha-Joon Chang quote in the epigraph: social phenomena are not inherently complex, and social explanations do not need to be complex to be accurate.

We can motivate our argument by examining projected changes in the Taiwanese population, assuming that fertility rates remain at contemporary levels. In order to illustrate the effect of population momentum, we focus on the population size of childbearing-aged women, i.e. women aged 15-50. The 2019 UN World

Population Projections forecast that with a constant fertility rate, the population of women aged 15-50 will decline from 6.672 million in 2020 to 2.208 million in 2100. We can use these numbers to calculate the factor by which this age group changes across successive generations:

$$6672 \cdot r^{80/35} = 2208$$
$$\Leftrightarrow \qquad r = 0.6165.$$

This means that the size of successive cohorts of women aged 15-50 will change by a factor of 0.6165, i.e. decline by about a third. Note that built into this number are assumptions about mortality and migration made by the UN, and cannot be used to calculate a strict fertility rate; however, multiplying the number by 2 (to give 1.233) is a rough estimate of the cohort fertility rate. Here we term the number 1.233 as the "generational replacement rate," i.e. the number of people in the next generation of men and women that will replace a couple in the present generation (assuming a constant sex ratio of 0.5). We can use this number to project forward to 2300, as shown in Table 1.1. Note that population projections to 2300 are not rare in demography (e.g. Basten et al. 2013; Population Council 2004); moreover, similar long-run climate projections (e.g. Horton et al. 2020) are frequently used to motivate concerns about what kind of world will be inherited by our descendants (e.g. Abnett 2021). Table 1.1 shows that by 2300, the number of women aged 15-50 will have declined to 2% of its present population.

Table 1.1: Projected numbers of women aged 15-50 in Taiwan 2020-2300, assuming a constant generational replacement rate.

Year	**Number of generations**	**Population size (thousands)**	**Percentage of women aged 15-50 in 2020**
2020	0	6,672	100
2055	1	4,113	62
2090	2	2,536	38
2125	3	1,563	23
2160	4	964	14
2195	5	594	9
2230	6	366	5
2265	7	226	3
2300	8	139	2

Notes: This projection assumes a constant generational replacement rate of 1.233, and an equal sex ratio at birth. The size of the population aged 15-50 therefore changes by a factor of 1.233/2 across successive generations.

Source: Own calculations, using figures taken or calculated from the projections of UN 2019.

Reher 2007 and others argue that population decline may not be a negative experience for anyone living during such a period, and that a general global population decline may even be desirable. But even if that were true, the question remains: when *should* fertility increase? If not in 2020, then in 2090? Or 2300? Certainly before 2657, when there will be no women aged 15-50 left. Additionally, there is reason to believe that the lives of those living into the next century would not be particularly positive, given that they would live in the ruins of a built environment designed to accommodate three times as many people. The anthropological "Emptiness" project (https://emptiness.eu/) aims to measure and understand the lives and experiences of the remaining populations in the depopulating rural areas of post-socialist Eastern Europe. As suggested by its name and associated research papers (e.g. Dzenovska 2011, 2018, 2020), the experience of living in these places is characterised by a distinct lack of expectation of a positive future.

As a final note on why low fertility itself is a problem, it is important to recognise the uniqueness of post-transitional fertility in human history (Reher 2007). Prior to the demographic transition – and most likely stretching back to the start of human populations – average fertility rates remained at a high level. The equilibrium between births and deaths ensured that populations could persist for millions of years. In the (so far) very brief moment of the post-transitional era, it is unclear whether or how a similar equilibrium can be established, and what social systems might evolve to sustain that equilibrium. The uniqueness of our historical moment is part of what makes persistent below-replacement fertility a problem: no society in history as has sustained a low vital-rate equilibrium for more than a few decades, and from that perspective it seems important to treat long-term fertility as a serious issue.

To summarise, we have argued that low fertility is a problem in and of itself, because sooner or later it will have to increase in order for the population to persist. It remains true that much of the discourse around low fertility has become toxic and recriminatory, that many of the putative short- and medium-term negative consequences may be overstated, and that it is important to conceptualise low fertility as the outcome of institutional malfunctioning. It may also be true that slight population decline in the short term will be a positive experience for those living through it, and that the contemporary panic about low fertility is more harmful than useful. However, it is also imperative to recognise that aggregate fertility levels cannot persist at a low level in the long run.

1.3 Addressing low fertility through public policy

In order to increase fertility from low levels, governments in Taiwan and elsewhere have turned to family policies such as baby bonuses, parental leave, and subsidised childcare (Caldwell et al. 2002; McDonald 2002). However, the role of such policy – as well as the descriptor "pro-natal" that is often used to describe them – remains contested by feminist scholars. In Taiwan, some critics have described such policies as a "sort of instrumentalization of women's body, taking women's body as an instrument to fulfill nation's target goal" (Lee 2009). Here we explore feminist critiques of pro-natal policy, and give a precise working definition of "pro-natal policies" that will be used throughout the remainder of the thesis.

At the 1994 International Conference on Population and Development in Cairo, governments and other agencies reached a consensus on the appropriate balance between sexual and reproductive health, anti-natal programs, and gender equality (PDR 2019). The consensus specified that government population policies needed to be formulated in such a way as to protect the rights of women. Despite this consensus, critics have characterised pro-natal policies in some East Asian countries as intrusive interventions into the private lives of women (e.g. Hiroko 2005; Sun 2012). Moreover, some pro-natal policies in countries such as Singapore have been couched in implicitly racist terms by government authorities, who seek to promote births among high-income and well-educated Singaporeans and reduce births among immigrants (Song et al. 2013). In South Korea, members of the 4B feminist movement have even go so far as to not only explicitly reject the pro-natal stance of the state, but also romance and sexual relationships (Lee and Jeong 2021). Hur 2013 argues that the South Korean state has undertaken a "long, deep, and continuous mobilization of women's biological reproduction," and calls for the deconstruction of developmental hegemony as it applies to women's reproduction. In this way, the existence of formal reproductive rights can be seen as insufficient for allowing truly free reproductive decision-making. Pro-natal policy interventions in East Asia can therefore be criticised both in terms of possible violations of sexual and reproductive rights, and in terms of not allowing for uninfluenced childbearing decisions.

In light of feminist criticisms of pro-natalism, we aim to arrive at a definition of this term which can allow for freedom in reproductive decision-making. We therefore define "pro-natal policies" as government policies which aim to increase fertility rates by reducing the costs of various aspects of childbearing and childrearing. These costs could be direct costs – such as payments for childcare –

or indirect costs, such as the burden of domestic labour shouldered by mothers (Gietel-Basten 2018). Under this definition, pro-natal policies have a large overlap with family policies. However, there are policies that could be considered a pro-natal policy but not a family policy, such as an attempt to promote egalitarian domestic gender norms by changing the curriculum (e.g. Shih and Wang 2021). The looseness of our definition hangs on the word "aim," in that singular policies (such as extending parental leave) will usually have multiple purposes and be drawn up by a range of groups with different interests. As such, many different policies relating to families, employment, and education can be seen through a pro-natal lens. However, unless otherwise stated, we will take "pro-natal policies" to be limited to the expansion of various entitlements through family policies. For an indicative range of family policies, see the OECD family database: https://www.oecd.org/els/family/database.htm#public_policy.

Given the discussion in Section 1.2 that low fertility is a problem in and of itself, we believe that the Taiwanese government ought to be mindful of the long-term consequences of low fertility. However, if reproduction is to be conceptualised in terms of societies rather than individuals, then there are two consequences for policy. Firstly, if childbearing is a *social* good, then the cost of childrearing – potentially the entire cost – should be borne by the whole of society (Burggraf 1993; Li et al. 2019). Secondly, pro-natal policy efforts need to account for heterogeneous preferences across the population. It would be an inverted ecological fallacy to assume that any specific individual or group of people need to have children themselves – the only thing that matters in the long run is whether *society* reproduces, not whether individuals reproduce.

Beyond these issues, the evidence of the efficacy of pro-natal policies is unclear (e.g. Gauthier 2007; Rindfuss et al. 2016; Stropnik and Šircelj 2008). It is not known which policies have positive impacts on fertility, and if so, which policies. Moreover, evaluations of pro-natal policies face idiosyncratic methodological problems, largely due to the fact that incentives supplied by those policies (e.g. parental leave) are only provided after the outcome of interest (i.e. having a child) occurs (Neyer and Andersson 2008).

1.4 Gaps in the literature and thesis aims

Low fertility in Taiwan has largely been explained in terms of social institutions and gender, though there is a lack of empirical support for this theory. Taiwan experienced rapid social and economic development since 1949, but traditional attitudes about familial gender roles persist in large parts of the population (McDonald 2000a; Raymo et al. 2015; Tu et al. 2017). Women in Taiwan experi-

ence gender-equitable social institutions in education and employment, but also face gender-inequitable domestic expectations after marriage. Aggregate fertility is theorised to have declined as a consequence, since increasing proportions of women are unable or unwilling to undertake the dual burden of career ambitions and domestic obligations (McDonald 2013). The role of gender inequity has shown to be a cause of low fertility in other countries (e.g. Cooke 2009; Dommermuth et al. 2017; Nagase and Brinton 2017), but there is currently no evidence confirming this theory in Taiwan. We therefore aim to evaluate the validity of gender equity theory in Taiwan.

One government policy that can alter the gendered distribution of formal and domestic labour – and therefore increase fertility – is parental leave, though there is a lack of evidence on this connection. A central problem with evaluating the causal impact of leave policies (as well as other pro-natal policies) is that incentives provided by those policies are provided *after* the event of interest occurs. This makes it difficult to evaluate leave policies using the standard statistical methods for causal inference in the social sciences (e.g. Shadish et al. 2002). Empirical evaluations of the impact of parental leave on fertility generally do not explicitly state the precise effects they are able to identify, nor do they attempt to interpret the meaning of those effects in terms of the policy goal of pro-natalism (Ang 2015; Dahl et al. 2013; Lappegård 2010). Therefore it is currently not possible to meaningfully synthesise the available evidence of the effect of leave on fertility, without first building an appropriate conceptual and methodological framework. We therefore aim to investigate whether it is possible to develop such a framework, and whether this framework reveals anything about the effect of leave on fertility.

In Taiwan the government has been pursuing a pro-natal policy agenda since the mid-2000s, but it is not known whether this has had any impact on fertility. The most costly policy implemented over this period has been the Parental Leave Allowance (PLA), introduced in 2009, which entitles parents to 6 months paid leave at 60% of their salary. We therefore aim to evaluate whether the PLA had any effect on fertility, and if so, for whom.

In terms of analytical scope, this thesis focuses on fertility within marriages, mostly because most childbearing continues to occur within the context of marriage (as discussed in the Background chapter) and partly because some of the main social survey data does not contain information on extramarital childbearing. Some authors have argued that the low fertility rates experienced across East Asia are largely due to a retreat from marriage and declining nuptiality (e.g. Cheng 2014; Cheng and Yang 2021; Tai et al. 2019). It also remains true that children born to unmarried mothers are likely to stigmatised, and that mar-

riage continues to be seen as a prerequisite of fertility (Esteve et al. 2020; Lee et al. 2007; Yu and Liu 2014). Unmarried women are less likely to be targeted by pro-natal policy efforts, and the impact of this on fertility is unclear. However, it remains true that the vast majority of births are had by married women, and so in order to evaluate the determinants of childbearing the main sample of interest is married women. This narrows the scope of our research in that we can only paint a partial picture of Taiwanese childbearing; however, we are still able to capture over 95% of all childbearing in Taiwan.

1.5 Data sources and methods

Social survey data in Taiwan are generally provided by the Survey Research Data Archive of Academia Sinica (https://srda.sinica.edu.tw/index_en.php), and the data used to evaluate gender equity theory and the PLA are both sourced from there. In order to empirically evaluate gender equity theory, we need micro-level data that record the balance of domestic labour within a couple, as well as subsequent childbearing behaviour. The only data source meeting this criteria is the Panel Survey of Family Dynamics, which has been recording data on family behaviours since 1999. In order to evaluate the impact of the PLA, we need either panel or repeated cross-sectional data on married women's employment and childbearing, before and after 2009. The Women's Marriage, Employment and Fertility Surveys fulfil these criteria, as well as having large sample sizes (about 20,000 respondents per survey). In terms of methods, we use logistic regression models to predict birth events in both cases.

In our review of the literature on the effect of leave policies on fertility, we can use access to academic literature through university subscriptions. We use standardised systematic review methods, and particularly guidelines issued by the Campbell Collaboration (Campbell Collaboration 2019).

1.6 Structure of the thesis

We address our research aims in three research papers, which have been submitted to academic journals over the course of the PhD. As shorthand, we refer to these as Paper 1, Paper 2, and Paper 3. Paper 1 investigates whether gender equity is a valid explanation of low fertility in Taiwan, by evaluating the relationship between husband's share of housework and subsequent fertility. Paper 2 is a systematic review of the effect of parental leave on fertility, in which we build a conceptual and methodological framework to decompose different types of effects, as well as synthesise the best available evidence on that question. Paper 3 is an

empirical study of the effect of the PLA on the fertility of women eligible for benefits, using the framework proposed in Paper 2. Since each of these papers are stand-alone pieces in their own right, there is partial repetition of introductory and background material between them. Before our three papers is a background chapter, which discusses the Taiwanese context, academic research on low fertility and pro-natal policies both globally and in Taiwan, and states our research aims. After the three papers, our conclusion rehearses the key findings of our research, how they meet our research aims, as well as implications, limitations, and avenues for further research.

2

Background

2.1 Introduction

Trends in low fertility at the aggregate level in Taiwan have been well documented by government statistics; by contrast, our understanding of the causes of low fertility in Taiwan – and whether pro-natal policies are increasing fertility – remains patchy. A body of sociocultural theory has long identified a central cause of low fertility in Taiwan to be the incompatibility of women's demand for gender-equitable social participation with the persistence of traditionally gendered social expectations for women's home life. However, there is no direct evidence that this incompatibility has decreased fertility. If this theory is valid, it would imply that policies that can facilitate women's labour force participation and encourage men to adopt gender equitable domestic roles – policies such as parental leave – would be an appropriate pro-natal strategy for the Taiwanese government. Indeed, the government introduced a major parental leave policy in 2009. However, the wider literature claims that the evidence for the effect of leave policies on fertility is mixed, and that evaluations of pro-natal policies generally face conceptual and methodological problems. Therefore it is difficult to predict the effect of parental leave on fertility in Taiwan without a clear conceptual framework and method, and by evaluating what the evidence globally shows.

This chapter provides background information on various topics relating to fertility, parental leave, and Taiwan. The purpose of this chapter is to arrive at a set of research aims, aims which are addressed by the research papers in the subsequent three chapters. We arrive at our research aims through reviewing the academic literature, and by identifying any gaps in that literature.

We draw on four distinct sources of information in this chapter. The first

source is the academic research literature on low fertility and Taiwan. The second source is the mostly non-academic literature on aspects of the history, economy, politics, and society of Taiwan. The third source is statistical data on aggregate trends in Taiwan, such as the TFR and total population. This data is mostly provided by the Department of Household Registration, Ministry of the Interior. The fourth source is legal documentation on pro-natal policies in Taiwan, which is also provided by the Taiwanese government.

This chapter is structured as follows. Firstly we summarise the Taiwanese context, in terms of Taiwan's political and economic history, aggregate statistical indicators of fertility, and the evolution of pro-natal policy efforts. Secondly we review the academic literature on several topics: low fertility in general, low fertility in Taiwan, gender equity theory, the effect of pro-natal policies on fertility, and low fertility in East Asia. We then use the contextual information and literature review to establish gaps in our understanding of low fertility in Taiwan and the effect of pro-natal policy on fertility. We use these gaps to generate a set of research aims, and then outline how our three papers seek to meet these research aims.

2.2 The Taiwanese context

2.2.1 Political and economic history

The history of contemporary Taiwan can be dated to 1949, when, at the culmination of the Chinese Civil War, the government and army of the Republic of China (ROC) – controlled by the nationalist Kuomintang party (KMT) – ceded mainland China to the Chinese Communist Party (CCP) and retreated to the island of Taiwan (Metzler 2017). This division has been maintained to the present day, with mainland China being officially known as the People's Republic of China (PRC), and Taiwan being officially known as the ROC. After 1949, Taiwan under the KMT pursued a program of rapid industrialisation and socioeconomic development. Over the second half of the 20th century Taiwan was transformed from a peripheral agrarian society to a highly-educated, developed society with a pivotal role in the global economy, particularly in the production of semiconductors. The socioeconomic transformation of Taiwan is known as the "Taiwan Miracle," and has been compared to similar transformations in South Korea, Hong Kong and Singapore over the same period (Wade 2004).

Taiwan remained under martial law from 1949 to 1987, under the KMT. The ROC during this period has been characterised as "state capitalist" and right-wing authoritarian, meaning that the government worked closely with private

capital to promote the development of a domestic industrial base, geared towards exporting to Western markets (Brett 2009; Wade 2004). Civil freedoms and union activity were suppressed, although average incomes, living standards and levels of education rose dramatically. Democratic and constitutional reforms were initiated in the late 1980s and continued through the 1990s, with the first democratically elected president – Lee Teng-hui of the KMT – being elected in 1996.

Since the 1990s, Taiwanese politics has been dominated by two factions: the Pan-Blue coalition, headed by the KMT, and the Pan-Green coalition, headed by the Democratic Progressive Party (DPP) (Metzler 2017). The Pan-Blue coalition tends to favour eventual reunification with mainland China, and is more socially conservative. The Pan-Green coalition tends to favour complete independence from the mainland, and is more socially liberal. Control of the presidency and the unicameral Legislative Yuan has oscillated between these two coalitions since democratisation. The Executive Yuan, responsible for conducting government affairs, is headed by a Premier who is chosen directly by the President.

2.2.2 Social and demographic transformations

Taiwan's economic modernisation was accompanied by a complete transformation of Taiwanese society. Moreover, the developmental orientation of the state meant that the government intervened in various social processes in order to produce a highly educated and disciplined workforce (Esping-Andersen and Billari 2015). Key social changes include the proletarianisation of the rural peasantry, urbanisation, mass education, the widespread provision of healthcare, and the development of the middle class (Metzler 2017). Changes that took centuries in Europe and North America were compressed into 50 years in Taiwan. A typical Taiwanese person born in 1930 would have started life as an agricultural labourer under a local landlord, and received little to no education (Wickberg 1975). By contrast, their grandchild born in 1990 would live in city, receive a university education, and use smartphones and computers with processor chips manufactured in Taiwan.

The population grew from 10 million in 1958 to 24 million by 2019, as shown in Figure 2.1. Population growth was driven by the demographic transition in Taiwan, as shown by the imbalance between births and deaths shown in Figure 2.2. The contemporary age structure of the population is illustrated by the population pyramid in Figure 2.3. The narrowing towards the base of the pyramid indicates that cohorts aged 30-50 have not been having enough children to replace themselves. In 30 years there will be far fewer individuals of childbearing age, meaning that future population growth will have a negative momentum.

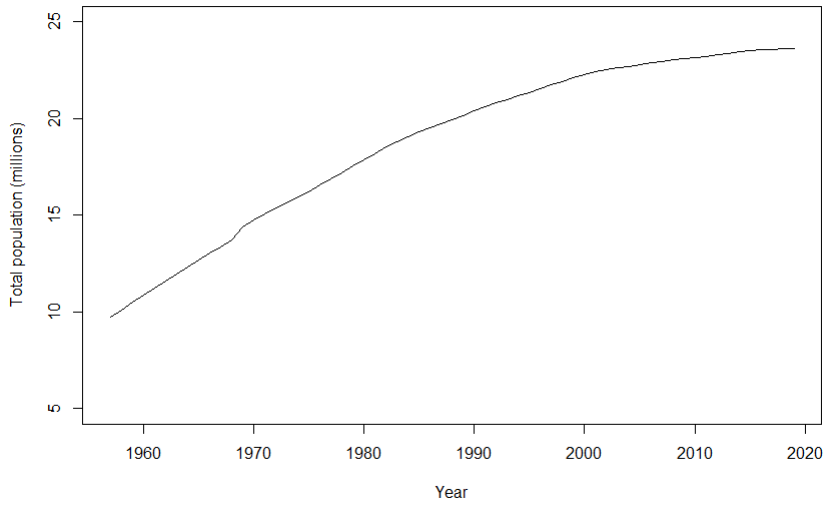


Figure 2.1: Total population of Taiwan, 1958-2019.

Source: Department of Household Registration, Ministry of the Interior (MOI). Available at https://www.ris.gov.tw/app/en/3911. Accessed on 02/01/2020.

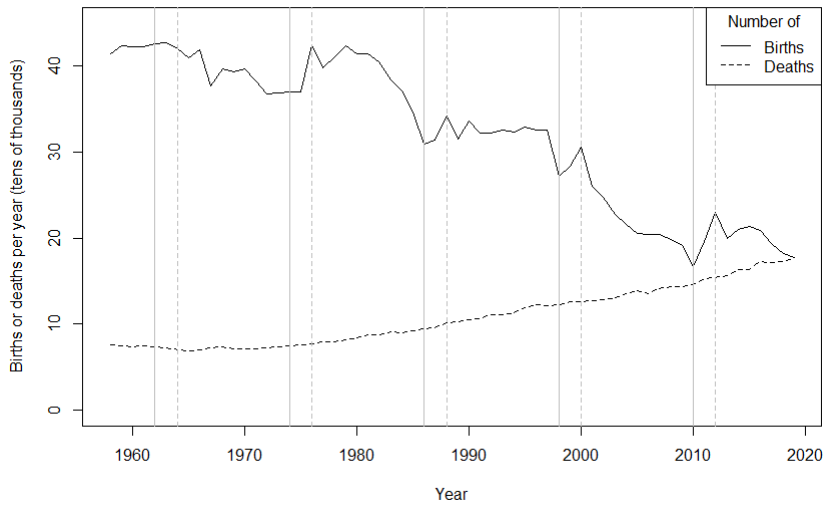


Figure 2.2: Numbers of births and deaths in Taiwan, 1958-2019.

Notes: Solid grey vertical lines indicate Years of the Tiger, and dashed grey vertical lines indicate Years of the Dragon.

Source: Department of Household Registration, Ministry of the Interior (MOI). Available at https://www.ris.gov.tw/app/en/3911. Accessed on 02/01/2020.

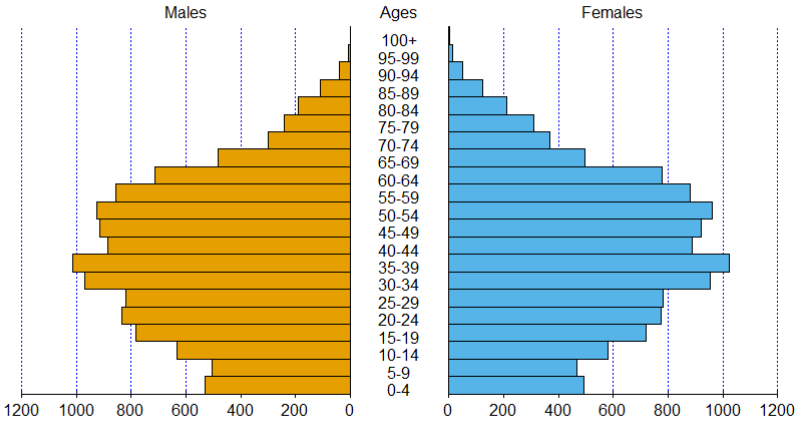


Figure 2.3: Population pyramid for Taiwan, 2015.

Notes: The population for each age category is in thousands. Graphic produced using the 'pyramid' R package (Nakazawa 2019).

Source: R package 'wpp2019,' containing estimates and projections from the UN World Population Prospects 2019 (UNPD 2020).

One important recent social change for fertility has been the 'massification' of higher education since the mid-90s (Chan and Lin 2015; Cheng and Jacob 2012). Following a policy of opening new universities and relaxing entry requirements, the percentage of high-school leavers going on to tertiary education increased from 57% in 1995, to 96% in 2015.[1]

2.2.3 Trends in fertility

The Total Fertility Rate (TFR) in Taiwan from 1949 to 2019 is illustrated in Figure 2.4. From a high of 7.04 births per woman in 1951, the TFR declined past the replacement level of 2.1 in 1984, indicating the completion of the demographic transition in Taiwan. The TFR remained roughly stable from the mid-80s to the late 90s, before resuming its decline in 1998. In 2003 the TFR passed the "lowest-low" threshold of 1.3 births per woman, and sank to just 0.895 births per woman in 2010. Since 2010 the decline in the TFR seems to have stopped, although it has remained below 1.3. In Figure 2.4 and in some subsequent Figures, the solid grey vertical lines indicate Years of the Tiger, and the dashed vertical lines indicate Years of the Dragon. In Taiwan it is considered unlucky to have a child in a Year of the Tiger, and lucky to have a child in a Year of the Dragon; numbers of births therefore tend to drop in Tiger years, and spike in Dragon years (Freedman et al.

[1] https://english.moe.gov.tw/cp-86-18943-e698b-1.html

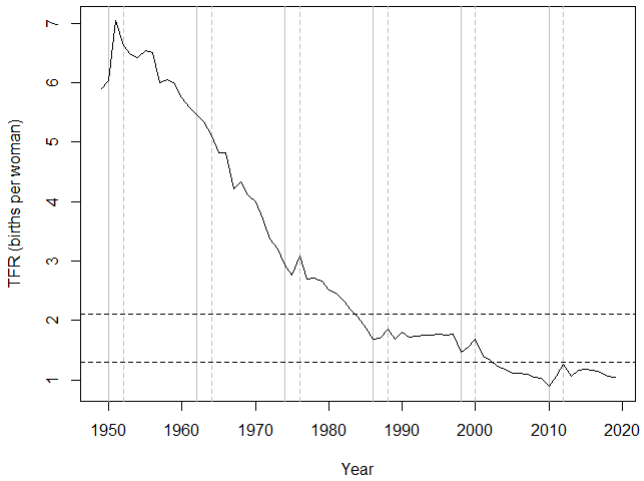


Figure 2.4: Total Fertility Rate, 1949-2019.

Notes: Solid grey vertical lines indicate Years of the Tiger, and dashed grey vertical lines indicate Years of the Dragon. The two dashed horizontal lines indicate TFR values of 2.1 (replacement level fertility) and 1.3 (threshold for lowest-low fertility).

Source: Department of Household Registration, Ministry of the Interior (MOI). Available at https://www.ris.gov.tw/app/en/3911. Accessed on 02/01/2020.

1994b; Goodkind 1991, 1993).

Trends in Age-Specific Fertility Rates (ASFRs) by 5-year age group are shown in Figure 2.5. From 1950 to the mid-70s the fertility of all women over 20 declined significantly, with the fertility of women in their 30s declining the sharpest. From the mid-70s to the present, fertility decline was driven by declining fertility for women in their 20s; the fertility of women in their 30s has broadly increased since the the mid-80s. Since 2009, ASFRs for women aged 30-39 have been higher than ASFRs for women aged 20-29. Women postponing their births from their 20s to their 30s in this period is also indicated by Figure 2.6, which shows trends in the average age of women giving birth and the average age of women having their first birth. Figure 2.5 also shows that fertility rates for women in their 40s have remained negligible since 1980, and fertility for women aged 15-19 declined steadily across the entire period.

Overall these trends indicate that post-transitional fertility in Taiwan has been characterised by women postponing their childbearing from their 20s to their 30s. However, increases in fertility at older ages has been smaller than declines in younger ages, indicating reduced fertility quantum.

2.2.4 Marriage and fertility

The vast majority of births in Taiwan are had by married women. Figure 2.7 shows the trend in the proportion of births that are given by married women, out

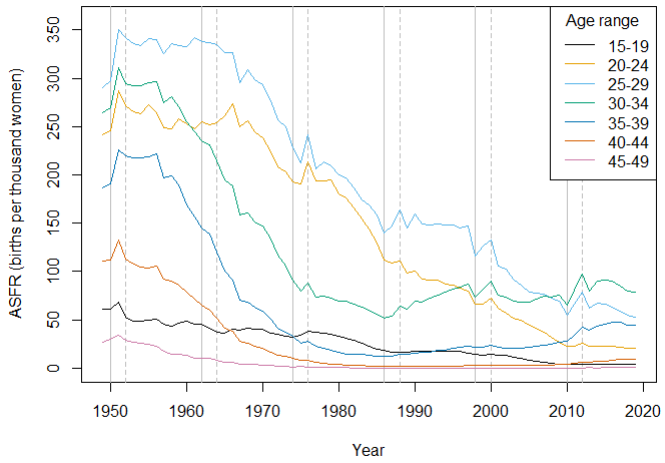


Figure 2.5: Age-Specific Fertility Rates, 1949-2019.

Notes: Solid grey vertical lines indicate Years of the Tiger, and dashed grey vertical lines indicate Years of the Dragon.

Source: Department of Household Registration, Ministry of the Interior (MOI). Available at https://www.ris.gov.tw/app/en/3911. Accessed on 02/01/2020.

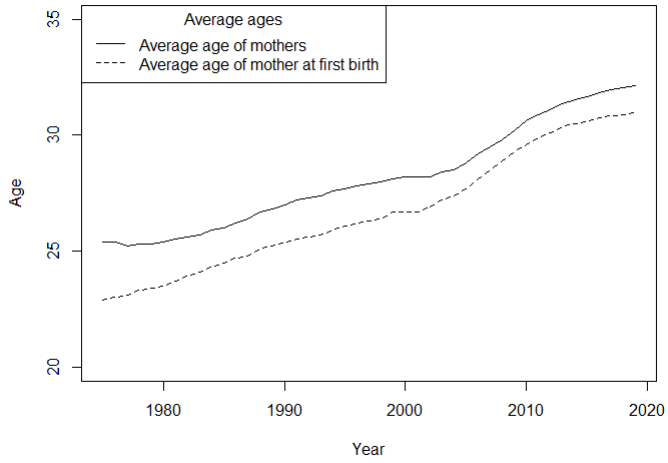


Figure 2.6: Average Ages of Mothers, 1975-2019.

Notes: The "Average age of mothers" trend line refers to the average age of all women giving birth in that year. The "Average age of mother at first birth" trend line refers to the average age of all women having their first birth in that year.

Source: Department of Household Registration, Ministry of the Interior (MOI). Available at https://www.ris.gov.tw/app/en/3911. Accessed on 02/01/2020.

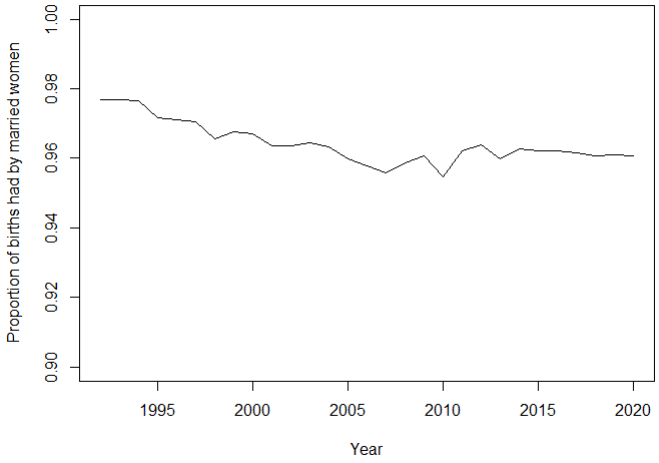


Figure 2.7: Proportion of all births had by married women, 1992-2020.

Source: Ministry of the Interior (MOI). Available at https://www.moi.gov.tw/stat/english/node.aspx?sn=7132. Accessed on 02/01/2020.

of all births. Since at least 1992 this proportion has remained between 0.95 and 0.98, and stood at 0.96 in 2020. Marriage is therefore a *de facto* prerequisite for childbearing in Taiwan, and so trends in marriage are likely to be important for understanding fertility.

The proportion of women who have ever been married by each age has been consistently declining across successive birth cohorts. This phenomenon is illustrated in Figure 2.8, which shows that among women aged 60-64 in 2016, 70% were married by the age of 25; by contrast, only 11% of women aged 25-29 were married by 25. Though younger cohorts do appear to still have large proportions getting married between 25-35 (indicated by the steep gradients between the ages of 25 and 30), the proportion ever married by all ages is lower in each successive cohort, suggesting that lower proportions of women in younger cohorts will have ever been married by age 50. Furthermore, the mean age at first marriage has been increasing steadily over time, as shown in Figure 2.9. The mean age at first marriage increased from 22 in 1971 to 30 in 2016 for women, and from 28 to 33 for men over the same time interval.

According to Yip and Chen (2016), fertility decline pre-1983 was primarily driven by declining marital fertility, while fertility decline after 1983 was largely due to declining nuptiality rates. Yip and Chen's conclusion is echoed by Chang (2006), who finds that after 1980 almost all of the decline in the crude birth rate (CBR) was accounted for by declining nuptiality. Decomposition analysis reveals that fertility decline between 1995 and 2005 for women aged 20-24 was driven entirely by falling nuptiality, and that marital fertility in fact increased by 44% (Lee 2009). Both marital fertility and nuptiality continued to decline for women

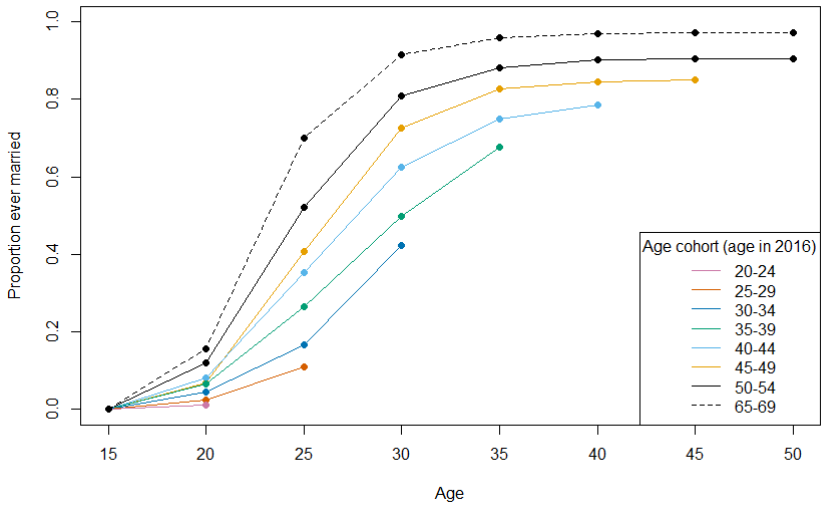


Figure 2.8: Proportion of women ever married by different ages, by age cohort in 2016.

Notes: Each curve represents a different birth cohort; for example, the solid black line represents women aged 50-54 in 2016. For a given birth cohort, a data point indicates the proportion of women that had ever been married by that age. For example, about 50% of women aged 50-54 in 2016 had been married by the age of 25. Data for this plot are from a sample survey, and are therefore not necessarily representative of the entire population. N=16,669.

Source: Women's Marriage, Employment and Fertility Survey (WMFES) 2016. Own calculations. Data available via application from the Survey Research Data Archive, Academia Sinica: https://srda.sinica.edu.tw/index_en.php.

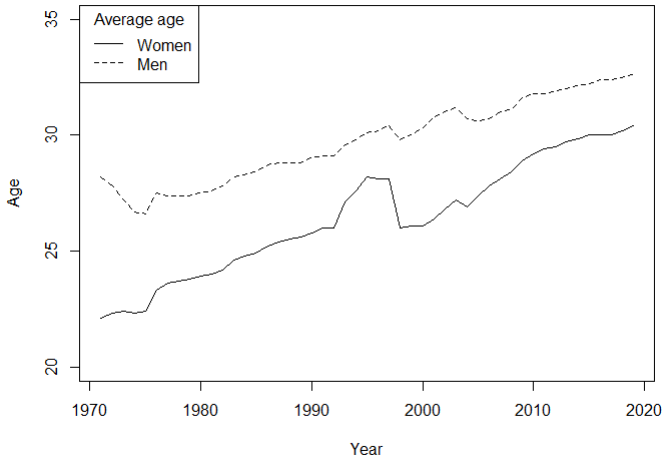


Figure 2.9: Mean age at first marriage for men and women, 1971-2019.

Source: Department of Household Registration, Ministry of the Interior (MOI). Available at https://www.ris.gov.tw/app/en/3911. Accessed on 02/01/2020.

aged 25-29 in 1995-2005, reversing the slight increase in marital fertility observed in 1985-1995.

Overall, these trends in marriage and fertility point to a society in which women are marrying later and less, postponing childbirth from their twenties to their thirties, and having smaller families (Chen and Chen 2014; Jones 2007). Low fertility can be understood in terms of declining marriage, since childbearing in Taiwan still occurs almost totally within marriages. While there is some evidence that marital fertility has increased for some age groups since the mid-90s, this has been insufficient to prevent a fall in the TFR.

2.2.5 Pro-natal policies

The Taiwanese government has increasingly sought to increase fertility since 2006 – the key government documents describing pro-natal policy efforts are the ROC Population Policy Guidelines, the Population Policy White Paper, and the Population Policy Data Collection (MOI 2013, 2014, 2020).[2] For a timeline of the development of pro-natal policy up to 2014, see Lee and Lin 2016.

In 2006, the national Guideline for Population Policy was revised to make pro-natalism a part of government policy; in the same year, the government also unveiled the Mega Warmth Social Welfare Program (MWSWP), which included provisions for extending maternity leave, providing paternity leave, childcare subsidies, and the provision of early childhood education and care (ECEC) (Chen 2012; Lee 2009; Lin and Yang 2009; MOI 2014). However, these new policies did not immediately translate into the provision of extra money and entitlements for parents: for example, the new Parental Leave Allowance (PLA) only became funded and available to parents in 2009, and the provision of ECEC had to wait for the passing of the ECEC Act in 2010 (Tsai 2012). In 2008 the government introduced the Population White Paper, which specified further policies for increasing fertility, as well as those in the Guideline for Population Policy (Lee and Lin 2016; MOI 2013). The Population White Paper was revised in 2012 and 2013, with the 2013 version specifying a goal of 180,000 births per year over the following decade (MOI 2013). In 2018 the government approved the Countermeasures Against the Declining Birthrate in Taiwan, which was further amended in 2019 (MOI 2020). These countermeasures include the expansion of public pre-schools, and providing publicly-subsidised placed at private pre-schools (MOI 2020).

[2] All are available at the Department of Household Registration (Ministry of Interior) website: https://www.ris.gov.tw/app/en/676.

2.3 Literature review

In this section we review the academic literature on three topics: low fertility in developed countries generally, low fertility in Taiwan, and low fertility in East Asia. The purpose of this section is to summarise the limits of what is currently known about low fertility in Taiwan, contextualised in terms of the literature on fertility in developed societies and East Asia generally. This summary will enable us to identify gaps in the literature. In our review of the literature on low fertility internationally, we focus on the socioeconomic and gender determinants of fertility, such as employment status and the gendered distribution of household labour. We chose to focus on these factors because our review of low fertility in Taiwan indicates them to be prominent factors. Therefore our review in Section 2.3.1 anticipates our discussion of low fertility in Taiwan in Sections 2.3.2. We also review the international literature on the effects of pro-natal policies on fertility; similarly, we review this literature because we later show that there has been a lack of evidence of the effects of pro-natal policies on fertility in Taiwan.

2.3.1 Research on fertility in developed societies

2.3.1.1 Trends and conceptual frameworks

At the macro-level, contemporary fertility in many high- and middle-income countries is understood through a range of conceptual and theoretical frameworks (Balbo et al. 2013). Here we discuss two such frameworks: the Second Demographic Transition (SDT), and social institutions.

Researchers have broadly understood fertility in contemporary Western societies through the framework of the Second Demographic Transition (SDT) (Lesthaeghe and van de Kaa 1986; Lesthaeghe 2020). After fertility declined to replacement levels or below due to the first demographic transition, rising levels of cohabitation, increases in ages at first marriage, and rising levels of extra-marital childbearing became distinguishing features of fertility in developed societies (e.g. Lesthaeghe and Neels 2002; Ogden and Hall 2004; Raley and Raley 2001). These changes in childbearing patterns have been driven by the liberalisation of social values, particularly regarding extramarital cohabitation and childbearing (Lesthaeghe 2020). Evidence for the SDT in Western societies has been broadly supportive (e.g. Ogden and Hall 2004). However, evidence supporting the SDT in East Asia is more mixed – though individuals are marrying later and less, cohabitation and extra-marital childbearing have not increased significantly (Lesthaeghe 2010; Matsuda 2019; Ochiai 2011; Ravaneral et al. 1999; Raymo et al. 2015).

A major development in global fertility since 2000 has been the bifurcation of fertility rates between Northern and Western Europe, and the Anglosphere (higher fertility), and Southern, Central and Eastern Europe, and East Asia (lower fertility) (Rindfuss et al. 2016). In 2012, the first set of countries had an average TFR of 1.9, whereas the second set of countries had an average TFR of 1.3 (Rindfuss et al. 2016). Theorists have attributed this divergence to differences in institutional arrangements and policies between the two groups, despite within-group heterogeneities in terms of institutional arrangements and policies (Rindfuss and Choe 2016; Rindfuss et al. 2016, 2003). Institutional arrangements in this context refers to social norms in education, labour, family, and gender relations, and policies to government laws and regulations concerning the family (Esping-Andersen 1990). Overall, countries with more traditionally gendered institutions appear to have lower fertility (Brinton and Lee 2016; McDonald 2000a; Rindfuss and Choe 2016).

2.3.1.2 Determinants of fertility

Many factors have been identified as determinants of fertility in developed societies; for reviews, see Balbo et al. (2013) and Schleutker (2014). Here we focus on microeconomic and institutional factors, and in the next section we discuss the role of gender equity.

A range of microeconomic factors have been identified as contributors to fertility decline and its persistence in modern societies, including income, education, and employment (Balbo et al. 2013). Income level can restrict fertility when parents seek to raise higher-quality (i.e. more expensive) children, and childbearing represents an income loss for women, which acts as a disincentive (Becker 1973, 1974; Becker et al. 1990; Kravdal 1992). Relatedly, women's education and employment aspirations increase the opportunity cost of having children, with highly educated women likely to postpone or forgo childbearing due to the high opportunity cost that this may represent for their future career development. On the other hand, highly educated women are more likely to partner with highly educated men (Oppenheimer 1994), which may reduce the household economic cost of having a child. Moreover, women who experience job displacement have been shown to have fewer children (Del Bono et al. 2015). Empirical evidence indicates that highly educated women at parity 1 or higher are likelier to have a child (Mencarini and Tanturri 2006; Mills et al. 2008; Wood et al. 2020). Also, highly educated women are likely to achieve comparable earnings to their partners (Behrman and Rosenzweig 2002), affording them bargaining power in relation to a more equal division of domestic labour. Together, these factors describe a situation in which women seek to balance their childbearing aspirations against

competing income and career demands, and particularly for professionals (Begall and Mills 2012; Miller 2011; Van Bavel 2010). In Western countries, a swathe of recent studies have demonstrated women's career planning to be a central factor explaining birth postponement (Amuedo-Dorantes and Kimmel 2005; Gustafsson 2005; Kneale and Joshi 2008; Miller 2011; Rondinelli et al. 2010). However, a meta-analysis of studies on the effect of women's employment status on fertility found institutional arrangements to be an important factor in explaining variations in effects across different societies (Matysiak and Vignoli 2008).

Institutional factors have become increasingly prominent in explanations of low fertility in developed societies – for an recent collection of essays on this topic, see Rindfuss et al. (2016). "Institutions" in this literature has at least four distinct (though related) meanings: social norms and expectations (such as the expectation that married women do the majority of housework); the availability of services and organisational arrangements (such as the availability of cheap childcare); prevailing macroeconomic conditions (such as the labour market being characterised by high unemployment); and family policies (such as the provision of parental leave). All meanings refer to factors which are external to particular individuals, couples and families, and which constitute the social environment in which childbearing decisions are made. Institutions relating to gender equity will be discussed in Section 2.3.1.3, and family policies in Section 2.3.1.4, and so here we indicate other institutional factors.

The literature on the effects of institutions focusses on factors such as the cost and availability of childcare (e.g. D'Albis et al. 2017; Wood and Neels 2019; Wood et al. 2016), female labour-force participation (e.g. Brinton and Lee 2016; Ma 2016; Matysiak and Vignoli 2008; Wood et al. 2016), and family support environments (e.g. Harknett et al. 2014; Pailhe et al. 2019; Yoon 2017). Each of these factors are highly country-specific, and so the particular configuration of institutional arrangements in each country are idiosyncratic; nonetheless, generalisations about which configurations promote fertility are possible (Rindfuss et al. 2016). Specifically, countries in which early-years childcare is widely available and socially acceptable, and in which mother's labour-force participation is facilitated by prevailing employment practices, are associated with higher levels of fertility. Ultimately, these institutional factors all concern the extent to which societies enable mothers to work and have children under terms equitable with fathers.

2.3.1.3 Gender equity and low fertility

As suggested by the preceding discussion on the role of institutions in determining contemporary fertility, over the past decade the academic literature has

become increasingly focussed on the importance of gender (in)equity in causing low fertility, and particularly in East Asia (e.g. Esping-Andersen and Billari 2015; Kato et al. 2018; Kim 2017; Li et al. 2020; Toulemon 2011). This literature is founded on "gender equity theory" as developed by McDonald (e.g. McDonald 2000a,b, 2009, 2013). Though researchers have proposed alternative theoretical frameworks – such as the "gender revolution" (Goldscheider et al. 2015; Raybould and Sear 2020) – we focus on McDonald's theory here for simplicity.

Gender equity theory argues that low fertility within a society is caused by incoherence between how different institutions regard the family. Institutions such as the labour market, education, family policies, and the family itself, implicitly assume men and women to have certain roles within the family. Taking the education system as a first example, in most developed societies it is socially expected that men and women participate in education on equal terms: both sexes can attend school and university, and there are no barriers to education uniquely placed on women. The education system therefore assumes a "gender equity" model of the family, in which gender does not determine a particular social or familial role (McDonald 2000a). As a second example of an institution, prevailing social norms about the family in some developed societies – such as in East Asia – dictate that families should ideally consist of husband who does not do housework or childcare, and a mother who does do housework and childcare. Therefore the social institution of the family assumes a traditional, "male breadwinner" model, in which domestic roles are highly gendered (McDonald 2000a). McDonald argues that in low fertility societies, education and labour markets tend to assume a gender equity model of the family, and that the institution of the family itself tends to assume a male breadwinner model (McDonald 2013). Women in low fertility societies therefore experience conflicting expectations about their roles in society: they are expected to receive an education and pursue a career, but also to stay at home and do all the housework and childcare. This incoherence makes it difficult for women to reconcile childbearing and work, and so fertility falls as a result. Gender equity theory implies that fertility in low fertility societies can be raised through the diffusion of gender egalitarian social norms (Esping-Andersen and Billari 2015), and that the promotion of such norms may be a viable pro-natal policy strategy in low fertility countries (Toulemon 2011).

Researchers have quantified gender equity in terms of gender roles and gender ideology, the division of domestic labour, and female labour force participation (e.g. Brinton and Lee 2016; Cooke 2009; Kato et al. 2018; Kim 2017; Li et al. 2020; Mills et al. 2008; Yoon 2016). Empirical studies have been mostly supportive of the validity of gender equity theory as an explanation of low fertility. A recent systematic review by Raybould and Sear (2020) synthesises 42 studies on the

impact of the gendered distribution of labour on realised fertility, and finds that more equal distributions of housework tend to be associated with higher fertility, both at the micro and macro levels. At the micro-level, domestic labour is most commonly quantified using one or more of three metrics: hours of housework, hours of childcare, and spouses' proportional share of these activities (e.g. Cooke 2009; Dommermuth et al. 2017; Nagase and Brinton 2017). The extent of the effect of household labour on fertility varies across countries, possibly due to policy measures aimed at increasing gender equity. In particular, increasing the provision of paternity leave could encourage fathers to take on a greater share of childcare and housework (Oláh 2003).

Micro-level empirical analyses of the effect of gender equity on realised fertility have tended to focus on second births, i.e. on having an additional child, for three key reasons. Firstly, second births are qualitatively different from births of first order, since two-child families are normative in modern societies, and so second births represent family building rather than family formation (Torr and Short 2004; Yoon 2016). Secondly, second and higher order births have declined as a proportion of total births, and this reduction is identified as a main driver of low fertility; therefore transitions to second and higher order births merit analysis for understanding trends in aggregate fertility (Cooke 2009; Goldscheider et al. 2013; Nagase and Brinton 2017). Thirdly, having a child introduces new forms of required domestic labour (e.g. childcare and child specific housework). The additional burden means the division of household labour for those with one child is likely to have a stronger effect on the likelihood of a next birth, compared to the division of household labour for those with no children (Cooke 2009; Nagase and Brinton 2017).

2.3.1.4 Efficacy of pro-natal policies

Pro-natal policies are government or local government family policies that can help individuals have children.[3] Family policies can lower the costs of financial childbearing and childrearing, enable parents to take time off work, or otherwise provide services (such as free school places for children above a certain age) that facilitate childcare. Family policies can be divided into child-related cash transfers, childcare subsidies, or financial support through the tax system (OECD 2019a). There is a broad empirical literature which attempts to evaluate the effects of specific pro-natal policies on fertility – for a non-systematic review,

[3]There are other, non-family policies that governments might pursue in order to encourage individuals to have children, such as encouraging marriages; for example, see the government of Singapore's promotion of marriage: https://www.msf.gov.sg/policies/Strong-and-Stable-Families/Pages/Promoting-Marriages.aspx. However, for our purposes we take pro-natal policies to refer exclusively to family policies.

see (Gauthier 2007). Overall, most commentators evaluate the evidence in this literature to be "mixed" (e.g. Bergsvik et al. 2020; Gauthier 2007; Hong and Sullivan 2016; Lappegård 2010; Matysiak and Szalma 2014), and so whether and how family policies impact fertility remain a matter of debate (Balbo et al. 2013; Hoem 2008; Olivetti and Petrongolo 2017). However, researchers have cautioned about certain conceptual, theoretical and methodological problems which may inhibit statistical identification of the causal effects of pro-natal policies (Gauthier 2008; Neyer and Andersson 2008; Ní Bhrolcháin and Dyson 2007).

Evaluating the fertility effects of pro-natal policies face certain idiosyncratic conceptual difficulties, which can lead to misleading policy implications if not properly addressed. In almost all cases, benefits provided by family policies are provided at or after the birth of a child – for example, individuals can only receive child benefits if they have already had a child. If we want to evaluate the effect of a policy on fertility, we therefore encounter a conceptual difficulty in that we are trying to identify the effect of something that is only realised through childbearing, on childbearing itself. By contrast, evaluating the effect of a medical intervention of subsequent health is much more straightforward: the intervention is given to some patients suffering from a illness, and then their health afterwards is recorded. This conceptual difficulty is not widely acknowledged in the literature evaluating the effects of pro-natal policies, though an exception is Lalive and Zweimüller (2009). There are some demographic papers that discuss theoretical, conceptual and methodological problems in evaluating the effects of pro-natal policies, such as Gauthier and Neyer and Andersson, as well as causal inference (Ní Bhrolcháin and Dyson 2007); however, none of these papers analyse the specific conceptual issue we raise. This conceptual issue matters because without carefully considering the precise effects being statistically identified within a given study design, researchers may reach misleading conclusions – and therefore give misleading advice to policymakers – about the effects of pro-natal policies. As a motivating example, Duvander et al. (2016) aim to evaluate the fertility effects of a parental leave reform in Sweden. Their study attempts to identify the effect of the policy by comparing the subsequent fertility on women giving birth shortly before the reform with women giving birth shortly after the reform. They have a strong study design which generates a clear difference between the two groups of women. However, the effect that their study design statistically identifies is the effect of having the new policy available for a child that a woman has had anyway, on a child she may decide to have at some time in the future. We contend that this effect is slightly dubious: generally speaking, when researchers and policymakers are interested in "the effect of a policy on fertility," what they really mean is the effect of the knowledge of having leave available for a theoretical future child,

on a woman's decision to have that child. There are many studies that identify effects in the same way as Duvander et al. (2016) (e.g. Cools et al. 2015; Farré and González 2019; Hart et al. 2019), which mostly find either no effect or very small effects. These papers collectively contribute to the conclusion in the literature that pro-natal policies may not impact fertility. However, that this conclusion has been reached without a clear demarcation of the different types of effects a policy might have on fertility, what study designs identify which effect, and what effects are of most importance for pro-natal purposes.

2.3.2 Research on fertility in Taiwan

This section explores research on fertility in Taiwan. In the first section, we explore research on the history of Taiwanese fertility, specifically the era of fertility decline and its immediate aftermath. Next, we provide an descriptive overview of the academic writing on Taiwanese fertility across a range of social scientific disciplines. In the following two sections we analyse parts of that literature in more detail, across two axes: gender roles in public and domestic life, and fertility intentions.

2.3.2.1 History of fertility in Taiwan

Having examined the literature on low fertility across developed societies generally, we now turn to the specific case of Taiwan. In this section we explore the literature on the history of Taiwan's fertility, as a prelude to a discussion of contemporary low fertility in the following section. A final section situates Taiwanese fertility within the regional context of East Asia.

There is a large body of literature which explores Taiwanese fertility through the period of fertility decline associated with the demographic transition, as well as post-transitional fertility before the onset of lowest-low fertility. As illustrated in Figure 2.4 transitional fertility decline began at some time between the late 1950s and early 1960s, with the TFR passing below 2 in the mid-1980s. For the purposes of this discussion, we will take the period 1960-1985 as representing transitional fertility decline, the period 1985-2003 as representing post-transitional stability, and 2003 to the present as lowest-low fertility.

Researchers have argued that the rapid fertility decline in the transitional period can be understood in terms of the successful family planning programmes and socioeconomic transformations engendered by the developmental state (e.g. Cernada et al. 2007; Freedman et al. 1994a). More broadly, transitional fertility decline has been understood as a component of the demographic transition in Taiwan (e.g. Freedman et al. 1964; Zhang 2011). Transitional fertility de-

cline was reported closely in the English-language academic literature as it happened, by a group of researchers including Ronald Freedman, Albert Hermalin, Te-Hsiung Sun, and Lolagene Coombs (Avery and Freedman 1970; Collver et al. 1967; Freedman 1965; Freedman et al. 1972; Freedman and Muller 1967; Freedman et al. 1963; Freedman and Takeshita 1964). This research programme was closely aligned with the developmental imperatives of the state, which aimed to reduce fertility in order to promote economic growth (Cernada et al. 2007).

A range of structural determinants of transitional fertility decline have been evaluated in the empirical literature, both during the transition and afterwards. For an early synthesis of these determinants, Hermalin 1974. Much of this literature – mostly published from 1970-1990 – focusses on the role of contraception and family planning, and sought to evaluate which specific family planning policy interventions were successful and potentially applicable to other high fertility settings (e.g. Chang et al. 1981; Chow 1965; Freedman et al. 1971; Montgomery and Casterline 1993; Sullivan et al. 1976; Sun et al. 1978; Sun and Ting 1989). In terms of structural changes in the society and economy, salient determinants of fertility decline include the fall in child mortality (Rutstein 1974), mass education (Freedman et al. 1977; Liu 1983), and changes in consumption, costs and benefits (Freedman 1975; Jejeebhoy 1978, 1979). Other factors such as secularisation (Schoonheim and Huelsken 2011) and agricultural reform (Hermalin and Lavely 1979) have been highlighted.

At the level of individuals and families, researchers have understood transitional fertility decline in terms of preferences for smaller families, declining birth intentions, and the influence of parental and extended families on childbearing decisions. Hermalin et al. 1979 and Freedman et al. 1975 show that declining birth desires and intentions strongly predicted lower rates of childbearing from 1967-1974 (see also Coombs 1979; Freedman et al. 1974; Sun et al. 1978; Yamanaka et al. 1982). However, though preferences for large families continued to decline into the early 1980s, fertility decline began to level off by that stage (Chang et al. 1981). Despite the weakening over the course of socioeconomic development, parental and familial ties persisted to some extent in influencing childbearing decisions (Chang et al. 1981; Coombs and Sun 1981; Yen et al. 1989).

Although the TFR remained stable in the post-transitional period of 1985 to 2003, fertility in Taiwan continued to experience various changes, as well as important continuities. As indicated by Figures 2.5, 2.6, and 2.9 in Section 2.2, both first marriage and childbearing continued to be postponed to older ages (Leete 1994). This period also saw the weakening of son preference and rise of gender indifference among Taiwanese couples (Lin 2009). Despite these changes,

strong family bonds and filial piety remained a feature of Taiwanese families, as illustrated by the significant remittances that adult children continued to send to their parents (Lee et al. 1994), and the persistence of grandparental co-residence (Chi and Hsin 1996).

Overall, transitional fertility decline in Taiwan can be seen as driven primarily through the developmental orientation of the state (Wade 2004). The state-directed development of the economy engendered rapid socioeconomic transformations which acted to reduce fertility, as did the provision of an extensive family planning programme aiming to reduce fertility. However, modernisation of the economy and public life were not accompanied by similar changes in the family, towards which traditional attitudes persisted beyond the transitional period.

2.3.2.2 Taiwanese fertility across social science

The literature on Taiwanese fertility can be roughly divided into several overlapping categories, which by and large follow disciplinary cleavages. Firstly, there are demographic papers that describe trends in fertility and nuptiality, as well as detailing changes in other potentially relevant quantitative variables such as ideal family size at different parities, mean age at first marriage, the divorce rate, and non-marital fertility rates. This set also includes papers that use statistical methods to explain demographic trends, that analyse trends in family planning, and that examine fertility in relation to mortality and migration. Secondly, there is the sociocultural anthropology and gender literature, which focuses on the changes and continuities in Taiwanese families, kinship, culture, and gender roles. Relatedly, there is a more nebulously grouped set of articles with a sociological or political-economic focus, which consider the effects of (amongst other things) economic modernisation, rising levels of education and aspiration, increased migration, and government policy, on fertility. Lastly, there are a set of articles that consider the microeconomic dimension of childbearing and childrearing, and have a particular emphasis on the cost of housing.

By far the largest of the four, the demographic subcategory has a much larger descriptive element than those of other disciplines, which documents various quantifiable aspects of Taiwanese fertility decline, including the age-distributional course of fertility decline (Freedman et al. 1994a), changes in parity-progression ratios (Feeney 1991), the fall in family size preferences and fertility intentions (Basten and Verropoulou 2015; Guttmacher 1975; Hermalin et al. 1979; Nair and Chow 1980), increases in ages at first marriage (Chang 1982; Wu 1970), changes in fertility due to immigration and cross-border marriages (Chang 1980; Chen 2008), and the spread of family planning (Chang 1982; Chang et al. 1981; Sun et al. 1978; Trewinnard 1998; Yen and Wang 1973). Articles that analyse the

determinants of fertility are divided between those that consider fertility per se (e.g. Lin and Yang 2009; Narayan 2006), those that cover marriage (e.g. Jones and Gubhaju 2009), and those that cover family planning. A segment of this literature is also dedicated to producing population projections for Taiwan (e.g. Kuo and Yue 2008).

An early notable article in the anthropological category is "Fertility as Mobility: Sinic Transitions," in which Susan Greenhalgh argues that fertility decline was driven by a combination of a continuing traditional economic culture within the family, and changing costs and benefits of children in the face of increasing education provision and declining infant mortality (Greenhalgh 1988). This theme of familial continuity and socio-economic change is picked up elsewhere, for example in Raymo et al. 2015, which discusses the continuation of "patriarchal, patrimonial, patrilineal, and patrilocal organization of East Asian families" in terms of Peter McDonald's gender equity theory of fertility, arguing that the disjunction between rapid socio-economic changes and familial and marital continuity have driven down fertility. This is because while economic and social opportunities for women have broadened significantly (due to factors such as access to better education), labour within marriage has remained highly gendered, which has led to a high opportunity cost for marriage and therefore low marriage and fertility rates. Other interesting aspects of fertility discussed in this article include the tightening of marriage markets (an absence of suitable husbands) due to a combination of improved female education and a traditional culture of female hypergamy (women marrying men of higher social standing), and whether rising rates of cohabitation and extramarital fertility could drive increases in the TFR (Raymo et al. 2015).

The sociological/political-economic literature is closely related to the anthropological literature, but has been separated from it here due to the latter's explicit focus on the family. In accordance with the wide body of theoretical and empirical research connecting improvements in female education with falling fertility in developing countries (Jejeebhoy 1995), much of this literature aims to specify this relationship for the Taiwanese experience (Anderson 1975; Cheng and Nwachukwu 1997; Freedman et al. 1977; Huang 2001; Liu 1983; Yang 2000). Another strand considers the effects of long-term socioeconomic development more broadly (Hsing 2003; Jia 1991; Poston 2000; Robey 1991; Sun 1984; Thangkasemvathana 1990), or considers the potential of further socioeconomic development to increase fertility (Chen and Liu 2007). Other topics included in this category consider the effects of changes in female employment (Chang et al. 1981; Chuang and Lin 2006; Jao and Li 2011; Li and Yang 2004; Stokes and Hsieh 1983; Yang 1981), the role of mass media and social networks in the diffusion of contraceptive

knowledge and fertility behaviour (Cheng 2011), and policy responses (potential or actual) to falling fertility (Freedman 1986; Frejka et al. 2010; Greenspan 1994). A notable recent paper in this last segment (which might equally be considered part of the microeconomic literature) considers the effects of education subsidies on fertility, arguing that were the government to extend the subsidies for public-sector workers to private-sector workers, the general fertility rate would increase by 20% (Keng and Sheu 2011).

Finally, the microeconomic literature evaluates the effect of changes in the costs and benefits of children on fertility. Topics covered in this category include the cost of housing and homeownership (Chen 2013; Lo 2012), unemployment and labour market structure (Cheng 2004; Huang 2003), the role of individual costs or benefits such as college tuition (Huang et al. 2006), financial support from children (Hermalin and Yang 2004; Lai 2011), and personal tax exemption (Huang 2002), as well as more general overviews of the economic dimension of childbearing (Liu and Hsu 2004; Mueller and Cohn 1977; Wang 1985).

2.3.2.3 Gender roles in public and domestic life

Theorists have attributed low fertility in Taiwan to a tension between family life and public life. In the family, traditionally gendered childcare and housework roles have persisted, while roles in education, labour, and politics have become much more gender equal (Frejka et al. 2010; McDonald 2009; Raymo et al. 2015; Tu et al. 2017). A key change has been the expansion of female higher education since the mid-1990s, reflecting a shift in women's opportunities for career progression (Cheng and Loichinger 2017), and a postponement of marriage and childbearing (Chen and Chen 2014). The share of 18-21 year-old women enrolled in tertiary education increased from 40% in 1995 to over 85% in 2006, and to 89% in 2017 (Chen et al. 2013; Chen 2016; MOE 2018). Evidence shows that increasing maternal education represents an important factor in contemporary fertility decline, by encouraging career aspirations for a greater proportion of women (Chen 2016).

In contrast, traditional family values and expectations have been slow in adapting to the changing aspirations of women in Taiwan (Raymo et al. 2015). Comparing gender-role attitudes, Yang (2016) finds that Taiwanese nationals are more likely than Chinese citizens to feel that women should remain at home, despite also expressing more gender-equitable attitudes towards education, political leadership, and economic leadership (Yang 2016). Furthermore, traditional values have been linked with upholding traditionally gendered domestic roles and behaviour even in an era of education and employment equality. Wives continue to bear most of the housework in Taiwan, which is explained by the hypothesis of

"gender display" (Hu and Kamo 2007; Kim 2013; Yu and Xie 2011). The gender display hypothesis argues that women who outperform their husbands in terms of their career seek to affirm their gender by doing more housework. In turn, doing more housework could dissuade women from having any or more children, since they would not want to add further to their already demanding domestic workload.

Together, these domestic continuities and broader contextual changes are argued to have led to persistently low fertility rates in Taiwan. Increasing career aspirations seem to have resulted in the postponement or renouncement of having children given a greater expected share of domestic workload on Taiwanese women, compared to men (Hori 2017; Qian and Sayer 2016).

2.3.2.4 Fertility intentions

Although aggregate fertility has stagnated at a low level since 2003, the evidence on fertility intentions, preferences and ideals presents a more nuanced picture. As a brief distinction between these variables: intentions correspond to planned future childbearing behaviour; ideals represent the number of children someone would like to have, absent any circumstantial barriers such as cost or physiological incapability; and preferences correspond to a set of childbearing outcomes, ordered in terms of how desired they are (Hin et al. 2011). Consequently, an individual's ideal family size is the same as their most preferred family size. However, these definitions and distinctions are not used consistently across the literature.

For a recent overview of fertility intentions, preferences and ideals in Taiwan, see Gietel-Basten 2018. In Taiwan, the two main data sources for fertility ideals and intentions come from the Taiwan Fertility and Family Surveys (TFFS) (derived from surveys initiated for the purposes of informing and evaluating Taiwan's family planning programme during fertility decline), and the Women's Marriage, Fertility and Employment Surveys (WMFES) (Gietel-Basten and Verropoulou 2018; Lee and Lin 2016). Using the TFFS data, Lee and Lin 2016 show that ideal family size has remained higher than actual fertility since the late 1990s, levelling out at 2.04 births per woman in 2012. This would seem to suggest that aggregate fertility could return to replacement level if barriers to childbearing were removed. However, the use of fertility ideals as a metric has been criticised by Basten and Verropoulou 2015, who argue that fertility intentions are a more reliable 'barometer' for childbearing attitudes. Their study decomposes fertility intentions for women at different parities, using WMFES data from 2006 and 2010. They find that most parity 0 women intend to have two children; however, a majority of parity 1 women do not intend to have another child, and 95% of parity 2 women do not intend to have another child. Supplementing the findings

of Basten and Verropoulou 2015, Freeman et al. 2018 conducted a series of in-depth interviews with childbearing-aged parents on their birth intentions. They find that parity 1 women who were expected to shoulder most of the childcare were reluctant to have a second child, due to those experiences. They also found that for women, gender inequity in the division of housework and childcare was seen as a barrier to leading a fulfilling social life. In this way, the findings of Freeman et al. 2018 would seem to explain why it is that women who initially intend to have two children revise that intention downward after the first child, as indicated by Basten and Verropoulou 2015. Returning to the WMFES data, Cheng and Hsu 2020 use the 2016 wave of that survey to evaluate the impact of the gendered distribution of housework and childcare on next birth intentions. They find that childcare equity is a significant factor in determining next birth intentions – and particularly for more highly educated women – but that there is no observable impact of housework equity on intentions.

One of the limitations of the empirical studies discussed above is that they all three restrict their analyses to married women of parity 1 or higher. This is explained by the fact that the WMFES (used by the latter two studies) did not ask unmarried women about their childbearing behaviour prior to 2016, partly because very few births in Taiwan are had outside of marriage (see Section 2.2.4). Nonetheless, we are unable to determine whether trends in the fertility ideals or intentions of unmarried women. For married women, the two-child norm is well characterised by the following quote from Gietel-Basten 2018 (p. 350): "Clearly, then, two children is as much a *ceiling* as, perhaps, an aspirational *ideal*."

2.3.3 Taiwanese fertility in the context of East Asia

Low fertility can be understood in terms of similar fertility patterns in other East Asian societies, specifically South Korea, Japan, China, Hong Kong and Singapore – for overviews, see Frejka et al. 2010; Gietel-Basten 2019; Jones 2007, 2019; Lin and Kamo 2015; Rindfuss and Choe 2016. These countries all share a Confucian cultural heritage and experienced rapid socioeconomic development in the post-war period, as well a rapid fertility decline from the 1950s (with the exception of Japan, where fertility decline began slightly earlier (Frejka et al. 2010)). Indeed, much of the literature covering Taiwanese fertility does so in a regional perspective. In this section we discuss regional trends and patterns in low fertility and its determinants, in order to highlight commonalities and differences between Taiwan and these other places.

As discussed in Section 2.3.1.1, TFR trends across economically advanced countries have bifurcated into two distinct groups: a higher group with a mean

TFR of about 1.9, and a lower group with a TFR of about 1.3 (Rindfuss et al. 2016). Singapore, South Korea, Japan and Taiwan all belong to the latter group, as would Hong Kong if it were included in the study (Frejka et al. 2010).[4] For cohorts of women born in the late 1960s, Total Cohort Fertility Rates (TCFRs) range between 1.2 and 2 across these five countries. Beyond fertility, marriage remains an important gateway to childbearing across East Asia, and trends towards later and less marriage have not been accompanied by concomitant increases in extramarital childbearing (Esteve et al. 2020; Jones and Gubhaju 2009; Jones 2007; Raymo et al. 2015). In contrast with other regions of Asia, mean ideal family sizes tend to be around 2 children or below (Casterline and Gietel-Basten 2018).

Many of the socioeconomic and institutional factors shaping Taiwanese fertility are also present across East Asia. In South Korea, Anderson and Kohler 2013 explain low fertility in terms of parental "education fever," i.e. high investments in children's education; a similar explanation for Taiwan is found in Lee 2009. As suggested by the importance of marriage for childbearing, traditional domestic gender roles have persisted to a large extent across the region (Raymo et al. 2015). The work of Kan and Hertog (Kan and Hertog 2017; Kan et al. 2019) demonstrates that unequal divisions of domestic labour between husbands and wives have persisted across China, Taiwan, South Korea and Japan, and that this reduces fertility preferences among women in all four countries. This echoes the conclusion of Suzuki 2013 (quoted in Gietel-Basten 2018), that low fertility levels in these latter three countries is explained by obstacles prohibiting childbearing, rather than innately low fertility ideals.

In a recent reflection on low fertility and policy responses across the six countries listed at the beginning of this section, Gietel-Basten 2019 argues that the problem of low fertility across the region is being badly conceptualised in the public discourse and by government policy. Gietel-Basten argues that low fertility in these countries – referred to collectively as "Low Fertility Pacific Asia" (LFPA) – should be understood as the *outcome* of "upstream institutional malfunctions," rather than a problem in and of itself. Echoing the empirical work discussed in Section 2.3.2.4, young people across the region still have ideals of childbearing, but existing institutional arrangements prevent them from doing so: "Few people in the region report at a young age an aspiration to be single and childless for their whole lives, but high percentages of people are" (p. 3) (Gietel-Basten 2019).

[4]Comparisons with the TFR in China is difficult because official estimates are unreliable (Gietel-Basten 2019; Guo et al. 2019). The UN estimated China's TFR as 1.665 in 2015, whereas Guo et al. 2019 estimated it as 1.047 based on a 1% sample census.

2.4 Gaps in the literature

The literature reveals several empirical and theoretical gaps (both explicitly and implicitly), as well as places where there is a clear scope for further research to extend the discussion.

- While there are a number of empirical studies evaluating the impact of domestic gender equity on fertility intentions in Taiwan (see Section 2.3.2.4), there are currently no quantitative studies that assess the impact of domestic gender equity on realised fertility. Therefore there is no direct evidence confirming gender equity theory as an explanation for fertility behaviour in Taiwan.
- The state of knowledge regarding the effect of pro-natal policies on fertility is hazy. It is not known whether most pro-natal policies actually have a positive impact on fertility.
- There is confusion in the literature on how to properly identify the effects of policies on fertility. Having a child effects covariates and the use of pro-natal policies, and so it is difficult to unpick the causes of childbearing from the effects of childbearing, both conceptually and analytically.
- It is not known whether the various pro-natal policy efforts in Taiwan have had any impact on the fertility rate. The TFR has stopped declining although it has remained below 1.3.

2.5 Aims of the thesis

- To empirically confirm whether gender equity theory accurately explains low fertility behaviour in Taiwan
- To develop a conceptual and methodological framework for understanding how pro-natal policies impact fertility
- To thoroughly evaluate the evidence for whether leave affects fertility generally
- To evaluate whether pro-natal policies in Taiwan have had any impact on fertility

3

Paper 1 – Declining Fertility in Taiwan: The Long-Term Deterring Impact of Housework Imbalance

Fertility in Taiwan has been persistently low since 2003, yet this pattern has remained largely unexplained. We seek to assess the way in which the division of housework influences the probability of having an additional child, and the probability of desiring an additional child. We evaluate whether fertility is effected in the short-term or the long-term, and develop a new argument to explain why we expect fertility to be more effected in the long-term. We find a significant association for long-term impacts, but no significant associations for short-term impacts. Specifically, the odds of a couple with an equal division of housework having a child within five years are 20.6% higher than that of a couple with the mean division of housework. This finding is significant at the 1% level.

3.1 Introduction

According to the UN (UN 2019), Taiwan has the third lowest Total Fertility Rate (TFR) in the world, with a period TFR of 1.13 in 2018. The long-term social and economic consequences of lowest-low fertility rates are potentially severe. As an increasing share of the population survives to older ages, rises in the dependency ratio are expected to erode the local human capital base, accelerating population ageing and putting pressure on the health care and pension systems (Bloom et al. 2011). Persistent low fertility also builds negative momentum into population growth, reducing both the absolute number of women in future reproductive age ranges and future workforce (Feeney 2003). Low fertility thus represents a major global policy concern, with governments pursuing pro-natalist policies via cash transfers, parental leave, and subsidised childcare (Lee 2009; Lin and Yang 2009).

To understand low fertility levels in developed countries, a body of literature has emerged that emphasises the role of gender equity (Balbo et al. 2013; Goldscheider et al. 2015; McDonald 2013). Conceptually, incoherence in the role of gender in social and domestic spheres is argued to result in low fertility levels. While some social institutions such as the family still prescribe traditional gender roles, contemporary social institutions such as education and employment are considered to adopt an egalitarian view of gender, and thus a large proportion of women are expected to forgo marriage or childbearing in favour of a career (Goldscheider et al. 2015; McDonald 2000a). These ideas have conceptually been used to explain low fertility in Taiwan, but yet there is a lack of empirical evidence (Frejka et al. 2010; McDonald 2009; Raymo et al. 2015; Tu et al. 2017). While there is some evidence for a domestic gender equality effect on fertility desire, there is no evidence for an effect on realised fertility behaviour (Kan and Hertog 2017). In particular, we do not know whether more equal distributions of housework and childcare are associated with a greater probability of having another child, and the ways in which fertility desire and behaviour change over time based on concurrent and past experiences of housework and childcare.

Equal distributions of housework and childcare can impact fertility in the short- and long-term. In the short-term, domestic gender inequality may lower fertility through expectations about the near future. Based on current experiences of unequal housework distribution, wives may feel that the responsibility of taking care of children will not be shared equally, and they may judge this situation as unfair and incompatible with their educational and career ambitions. At the same time, domestic gender inequality may only affect fertility in the long-term. Expectations are slow to adjust to year-on-year changes in domestic labour distribution, as recent past experiences influence present human behaviour

(Cheung and Kim 2018).

In this study, we seek to statistically assess the influence of the division of housework on the probability of future birth occurrences, drawing on data from the Taiwanese Panel Survey of Family Dynamics (PSFD), from 2010 to 2012 (annually), 2014 and 2016. Specifically, we use panel data binary logistic regressions to estimate the strength of associations between the share of household work and the probability of having an additional child (i.e. realised fertility), and the intention to have another child (i.e. desired fertility) in the short- and long-term. We expect the division of housework to have a larger effect on the probability of having another child in the long-term than in the short-term. We present a new argument to explain why we expect the long-term effect to be larger than the short-term effect. Our argument posits that the divisions of housework between couples are autocorrelated over time, meaning that relatively gender equitable young couples will become relatively gender equitable old couples, and that relatively gender inequitable young couples will become relatively gender inequitable old couples. We motivate this part of our argument with a discussion of the empirical evidence on the interrelationships over time between housework, equity, and marital satisfaction. We then argue that, over time, the autocorrelation of divisions of housework between couples may cause gender equitable couples to be cumulatively more likely to have another birth, relative to gender inequitable couples. This argument implies that – after observing the divisions of housework for a sample of couples – the longer the follow-up period, the larger the effect of housework division on the probability of a birth.

The next section provides an exposition of gender equity theory and evaluates the evidence for gender equity theory, before going on to consider some key social changes impacting gender and family in Taiwan, including the expansion of higher education for women. This section then argues that gender equity theory – as originally formulated by (McDonald 2000a, 2013) – implies that the relationship between gender equity and fertility is likely to change over time. The following section explores how gender equity and fertility change as couples age over the life course and develops a new argument, in which we expand gender equity theory to incorporate the dynamic relationship between gender equity and fertility over time. The data section describes our source of data, the Panel Survey of Family Dynamics (PSFD), and discusses the restrictions we apply to obtain our analytical sample. We then specify three sets of models, which concern short-term realised fertility, fertility intentions, and long-term realised fertility. The penultimate section discusses our model results and their policy implications, and a final section presents our conclusions.

3.2 Conceptual framework

3.2.1 Gender equity theory

The effects of gender equity have been studied in the wider fertility literature through the lens of gender equity theory as developed by McDonald (2000, 2013). This theory suggests that low fertility is caused by incoherence between the models of the family assumed by different family-oriented social institutions. The male breadwinner model, in which the husband works and the wife takes care of the children, is contrasted with the gender equity model, in which gender does not determine which partner does which type of work (McDonald, 2000). Then, if social institutions in education and the labour market, which presume a gender equity model, coexist with traditional family values and expectations, which assume a breadwinner model, women are likely to postpone or forgo having a child given conflicting expectations on career and childbearing aspirations (McDonald 2013). Higher gender inequity will also cause women to be less likely to desire having a child, since they know that the labour associated with having a child (e.g. childcare, housework) will be shared inequitably.

Empirical evidence testing gender equity theory seems mixed. Prior work has shown that more unequal distributions of household labour reduce the likelihood of having an additional child across a number of Western societies and East Asian countries (Brinton and Lee 2016; Matthews 1999; Nagase and Brinton 2017; Torr and Short 2004). Torr and Short (2004) find that American couples in which women do less than 54% of housework are 253% more likely to have another birth within 5 years than couples in which women do 54-84% of the housework. In Germany, Cooke (2004) finds that husband' s percentage share of housework has no effect on the likelihood of a second birth, but that husband' s percentage share of childcare increases the likelihood by one percent per percentage point of childcare share (Cooke 2004). Similarly, Cooke (2009) finds that husband's percentage share of childcare increases the likelihood of a second birth by five percent per percentage point of childcare share in Italy, but finds no association in Spain (Cooke 2009). In Norway, Dommermuth et al. (2017) find no effect of the division of childcare on subsequent fertility for couples at any parity, but also that parity 2 couples with a semi-equal division of housework are 104% more likely to have a third child than couples in which wives do most of the housework (Dommermuth et al. 2017).

3.2.2 The Taiwanese context

In Taiwan, low fertility has been conceptualised as and attributed to a tension between family life and public life. In the family, traditionally gendered childcare and housework roles have persisted, while roles in education, labour, and politics have become much more gender equal (Frejka et al. 2010; McDonald 2009; Raymo et al. 2015; Tu et al. 2017). A key change has been the expansion of female higher education since the mid-1990s, reflecting a shift in women's opportunities for career progression (Cheng and Loichinger 2017), and a postponement of marriage and childbearing (Chen and Chen 2014). The share of 18-21 year-old women enrolled in tertiary education increased from 40% in 1995 to over 85% in 2006, and to 89% in 2017 (Chen et al. 2013; Chen 2016; MOE 2018). Evidence shows that increasing female education represents an important factor in contemporary fertility decline by encouraging career aspirations for a greater proportion of women (Chen 2016). In contrast, traditional family values and expectations have been slow in adapting to the changing aspirations of women in Taiwan (Raymo et al. 2015). Wives continue to bear most of the housework in Taiwan: estimates for the proportion of housework done by wives ranges from 72% to 81% (Hu and Kamo 2007; Kim 2013; Yu and Xie 2011). In a context of increasing career aspirations, the expectation of a heavy and unequal domestic workload seems to have resulted in the postponement or renouncement of having children by many Taiwanese women (Hori 2017; Qian and Sayer 2016).

Empirically, there is one study that assesses the effect of domestic labour balance on fertility in Taiwan. It shows that women whose husbands do more housework have a greater desire for more children (Kan and Hertog 2017). The study focusses on the cross-sectional impact of housework division on realised fertility. There are no empirical studies that assess the effect of domestic labour balance on subsequent fertility behaviour, or the effect of unequal domestic work balance on realised and desired fertility patterns over the life course.

Prior empirical analysis focuses on births of second order, i.e. on having an additional child, which is justified for three key reasons. Firstly, second births are qualitatively different from births of first order, since two-child families are normative in modern societies, and so second births represent family building rather than family formation (Torr and Short 2004; Yoon 2016). Secondly, the number of second and higher order births have declined and this reduction is identified as a main driver of low fertility in Taiwan (Gietel-Basten 2018), and so transitions to second and higher order births merit analysis for understanding trends in aggregate fertility (Cooke 2009; Goldscheider et al. 2013; Nagase and Brinton 2017). Thirdly, having a child introduces new forms of required domestic

labour (e.g. childcare and child specific housework). The additional burden means the division of household labour for those with one child is likely to have a stronger effect on the likelihood of a next birth, compared to the division of household labour for those with no children (Cooke 2009; Nagase and Brinton 2017).

In light of the preceding discussion, we hypothesise:

Hypothesis 1. *For married couples*[1] *at parity 1 or greater, husbands doing a greater share of housework is positively associated with the likelihood of having an additional child.*

Hypothesis 2. *For married couples at parity 1 or greater, husbands doing a greater share of housework is positively associated with the likelihood of desiring an additional child.*

3.2.3 Extending the gender equity framework

Gender equity theory is static. It explains how gender equity affects fertility at a single point in time. However, gender equity within a marriage evolves over time, as spouses age, progress in their careers, acquire new responsibilities and have children. The effect of temporal changes on the relationship between domestic gender equity and fertility remains theoretically unspecified and empirically unexplored. To address this gap, we firstly discuss some of the key ways in which domestic gender equity and fertility change over time. Secondly, we theoretically develop the links between domestic labour, equity and fertility and argue that the impact of gender equity on fertility is likely to be stronger over the long-term.

Domestic labour equity varies over time. Evidence suggests that men's share of housework decreases over the life course, and particularly after childbirth (Baxter et al. 2008; Grunow et al. 2012; Luo et al. 2018). Men's decreasing housework share after childbirth is partly explained by mothers taking maternal leave while fathers remain at work, and partly explained by the new kinds of housework required in looking after an infant (Baxter et al. 2008). In terms of fairness, women report higher perceptions of fairness in housework after the birth of a child, and a declining perception of fairness thereafter (Perales et al. 2015). Theorists have explained why women find housework balance fairer after childbirth in terms of the extra time at home available to mothers taking leave (Perales et al. 2015). If fathers maintain their working hours to support the family, then mothers tend to find the new post-birth division of housework fairer, even if their share increases (Perales et al. 2015).

[1]Childbearing in Taiwan still overwhelmingly occurs within marriages, which is why we restrict our Hypotheses to married couples (Raymo et al., 2015).

Domestic labour gender equity effects how happy partners are with their marriage overall, reflecting marital satisfaction. In terms of housework, a widely-cited article in the sociological literature states that "the link between the division of household labour and marital quality is fairly well established" (Frisco and Williams 2003), a finding which has been also established in East Asian contexts (Qian and Sayer 2016). Specifically, couples with more equal arrangements have higher marital satisfaction (Frisco and Williams 2003). As couples get older, marital satisfaction becomes less responsive to changes in housework (Cheung and Kim 2018). Additionally, earlier marital satisfaction is a good predictor of later marital satisfaction, although marital satisfaction for all couples decreases over time (Lupri and Frideres 1981; Roberson et al. 2018; VanLaningham et al. 2001). The autocorrelation of marital satisfaction over time implies that differences in marital satisfaction between couples – as influenced by housework balance at a younger age – is likely to persist through time. Since marital satisfaction is less responsive to housework for older couples, these differences between couples may persist regardless of changes in housework balance at older ages. In terms of fertility, evidence suggests that marital satisfaction is a positive predictor of fertility for both genders (Parr 2010). Moreover, women that are dissatisfied in how their work and family lives are reconciled are less likely to have a second child (Luppi 2016), and individuals that report higher levels of subjective well-being (i.e. happiness) are more likely to have children (Cetre et al. 2016; Le Moglie et al. 2015; Mencarini et al. 2018).

To summarise the key points so far: firstly, the division of household labour is closely related to marital quality, although changes in the division of household labour affects marital quality less over time; secondly, couples that had relatively high marital satisfaction at younger ages will maintain relatively high marital satisfaction at older ages, irrespective of changes in housework at older ages; and thirdly, couples with higher marital satisfaction have higher fertility. Altogether, these three points suggest that as couples get older, couples that had a more gender equitable division of labour at younger ages will have a higher probability of having another birth than couples that had a less gender equitable division. Moreover, these relative probabilities between gender equitable couples and gender inequitable couples should remain stable over time, *ceteris paribus*. Since the probability of another birth would be higher for the gender equitable couple in every time period as they age, this could lead to a cumulative effect in the relative probabilities of another birth between these two groups of couples. Consequently, couples that had a more equal housework balance when they were younger could become increasingly likely to have another child as time passes, compared with couples that had a less equal housework balance when they were younger.

While the effect of housework balance on fertility can be tested both over the short-term and the long-term, assessing the impact of housework on desired fertility is less meaningful. The conceptual links between current housework responsibilities, perceived housework and desires for having a child in the long-term are unclear and difficult to elaborate, as desires for children are more ephemeral and vary over the life course (Hayford 2009). For this reason, we only evaluate the long-term effect on fertility behaviour, and not on fertility desire.

Following the argument put forward in this section, we hypothesise that:

Hypothesis 3. *For married couples at parity 1 or greater, husbands doing a greater share of housework has a larger positive effect on the likelihood of having another child in the long-term than in the short-term.*

3.3 Data

We draw on data from the Taiwanese Panel Survey of Family Dynamics[2] (PSFD 2019), a nationally representative, longitudinal survey gathering data on household relationships and behaviours, including information on respondents' and their partners' age, housework, education, and income. Respondents are selected by a stratified three-stage sampling procedure using household registration data. The observational unit in the PSFD is individuals, and the PSFD gathers information on partners and spouses (if the respondent has a partner or spouse) by asking respondents, not by asking partners or spouses directly.

The PSFD comprises four cohorts, starting in 1999, 2001, 2003, and 2009, at age ranges 34-46, 45-65, 26-39, and 25-32, respectively. While these age ranges are for the respondent, the age of respondents' spouses can be outside of these age ranges. Respondents were annually followed up until 2012, and biennially after this year. We focus on the 2003 and 2009 cohorts, since women in those cohorts (either respondents or spouses of respondents) were mostly of childbearing age in 2010. We were not able to use the data from the year 2009, because data on fertility intentions were only collected in surveys after 2010. We therefore model from 2010.

The 2003 and 2009 cohorts had an initial response rate of 47.33% and 49.86% respectively, which is typical for longitudinal surveys in developed countries (Rindfuss et al. 2015). Drawing on Rindfuss et al.'s review (2015), we assume these low response rates do not significantly bias our inferences. In that review, the authors evaluate the impact of response rates in a longitudinal survey of Japanese families in the 2000s (with a response rate just over 50%). The au-

[2]https://psfd.sinica.edu.tw/web/plan_01en.htm

thors find that, while response rates vary with demographic and socioeconomic variables, there is no effect of non-response on the relationship between socioeconomic or demographic predictors and fertility. Because the PSFD data is also longitudinal, has a comparable response rate to the Japanese study, is also from East Asia, and covers the 2000s and 2010s, we assume that the relationships between predictors and fertility in the PSFD data are unaffected by their response rates. Under this assumption, the response rates of the 2003 and 2009 surveys do not risk biasing statistical associations between housework and fertility.

We further restricted the analysis to married heterosexual couples, with one or more children, who remained married until 2016, took part in all five follow-up surveys after 2009 (2010, 2011, 2012, 2014, and 2016), and in which the wife was between the ages of 20 and 40 in 2010. This selected sample represents a balanced panel of 544 couples (260 female respondent couples and 284 male respondent couples) over the four waves from 2010 to 2014, producing a total analytical sample of 2,176 person-years. Data from the 2016 wave was needed because our models that predicted future childbearing have a lagged dependent variable.

3.4 Methods

In this section, we first describe the modelling approach adopted, and then discuss its empirical implementation.

3.4.1 Models

In the literature, panel data have been used to test the effect of household labour on next births. Panel data enable the analysis of the impact of domestic labour on fertility behaviour by assessing the influence of domestic labour at time t on the probability of having an additional child at time $t+1$. Panel data also enable controlling for unobserved heterogeneity of respondents, such as respondents' gender role ideology. This mitigates estimation bias for outcome variables like subsequent fertility, as the effects of unobserved variables are omitted.

Nine separate models are estimated. Six separate logistic regression models are estimated to assess the short-term effect of housework balance (i.e. year-to-year changes in housework balance) on fertility: three models focus on realised fertility, and three models focus on desired fertility. Additionally, three models are estimated to assess the long-term impact of fertility imbalance on realised fertility outcomes. In all three sets of three models, one model uses data on all couples, one model uses data for which the wife is the respondent, and one model

uses data for which the husband is the respondent. We adopted this approach to account for potential gender bias in the survey. Respondent's gender may affect the reported hours of housework for each spouse, given that wives and husbands are likely to have different perceptions of their relative contributions to housework (Kiger and Riley 1996).

Three models (Models (1)-(3)) on the short-term effects of housework on realised fertility are formally represented by the following equation:

$$ln\left(\frac{Pr(y_{i,t+1}|\alpha_i, \boldsymbol{x}_{it}, \boldsymbol{z}_i)}{1 - Pr(y_{i,t+1}|\alpha_i, \boldsymbol{x}_{it}, \boldsymbol{z}_i)}\right) = \alpha_i + \boldsymbol{x}'_{it}\boldsymbol{\beta} + \boldsymbol{z}'_i\boldsymbol{\gamma}. \tag{3.1}$$

In Equation 3.1, $y_{i,t+1}$ is a binary variable that takes the value of 1 if the couple had a child in year $t+1$; α_i refers to the time-invariant, couple specific effect; and $\boldsymbol{x}_{it}$ is a vector of time-varying independent variables in year t. The $\boldsymbol{x}_{it}$ vector includes the husband's proportion of housework, as well as control variables such as income and age. The $\boldsymbol{z}_i$ term is a vector of time-invariant independent variables, such as education. All variables are specified in the next sub-section. Vector $\boldsymbol{\beta}$ contains coefficients for the time-varying independent variables, and $\boldsymbol{\gamma}$ contains coefficients for the time-invariant independent variables. The vector $\boldsymbol{\beta}$ contains our main coefficient of interest, husbands' proportion of housework. We expect our estimate of this coefficient to be positive, indicating that husbands' proportion of housework has a positive short-term effect on realised fertility.

Generally, fixed-effects or random-effects methods are used to estimate panel data models. Fixed-effects models use variation within the same observational unit over time to estimate coefficients, meaning that fixed-effects models cannot estimate the effects of time-invariant variables. Random-effects models use variation between all observations to estimate coefficients, but require that error terms be uncorrelated with the independent variables. In our case, it is likely that the requirement that error terms are uncorrelated with the independent variables is not met, as gender equity is not directly observed (the division of housework is a proxy for gender equity). Therefore some correlation is likely to exist between the error term and other explanatory variables in our models (Goldscheider et al. 2013). As a result, this may bias random-effects estimates.

We therefore use a hybrid method, which combines fixed- and random-effects approaches (Allison 2009; Schunck and Perales 2017; Wooldridge 2010). Hybrid methods produce two sets of coefficients for time-varying variables: (1) within estimates, which are derived from variation within individuals over time; and (2) between estimates, which are derived from variation between different individuals at the same point in time (Allison 2009). For time-invariant variables, only the between estimates are produced (Schunck and Perales 2017). In this way, hybrid

methods enable unbiased estimation of time-varying variables as in the fixed-effects case, while also enabling unbiased estimation of time-invariant variables. In this paper, within estimates indicate whether changes over time in individual family or work commitment circumstances affect fertility behaviour, and between estimates indicate whether differences between families affect fertility at a given point in time.

The same hybrid method is used to estimate three models on the short-term effect of housework on desired fertility (Models (4)-(6)), which are represented in the following equation:

$$ln\left(\frac{Pr(y_{it}|\alpha_i, \boldsymbol{x}_{it}, \boldsymbol{z}_i)}{1 - Pr(y_{it}|\alpha_i, \boldsymbol{x}_{it}, \boldsymbol{z}_i)}\right) = \alpha_i + \boldsymbol{x}'_{it}\boldsymbol{\beta} + \boldsymbol{z}'_i\boldsymbol{\gamma}. \tag{3.2}$$

In Equation 3.2, y_{it} is a binary variable that takes the value of 1 if the couple desired a child in year t, and α_i, $\boldsymbol{x}_{it}$, $\boldsymbol{z}_i$, $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ have the same meanings as in Equation 3.1. Note that in Equation 3.1, the dependent variable is taken from year $t + 1$, whereas in Equation 3.2, the dependent variable is taken from year t.

The long-term models are estimated using a model specification which does not use the panel structure of the data. This specification intends to capture the effect of housework on fertility over the longer term using independent variables measured in 2010 to explain the birth of a child between 2011 and 2015. By "long-term effect on fertility," we mean whether or not a couple has a birth at any time over the next five years; equivalently, whether or not a couple has had a birth by 2015. In this way, the long-term effect does not refer to a lagged effect of husband's proportion of housework on fertility: we do not mean that husband's proportion of housework will have a bigger impact on the probability of a birth in 2015, than in 2011.

The long-term models are required to test Hypothesis 3 – that housework division affects fertility more over the long-term than the short-term – by comparison with the short-term estimates from Models (1)-(3). Moreover, the long-term model will provide an additional test of Hypothesis 1 – that is, that housework division affects realised fertility – by estimating the long-term effect of husbands' share of housework on fertility.

We use the proportion of husband's housework in a given year to predict the probability of a birth over subsequent years (Dommermuth et al. 2017). This measure may have the drawback of representing domestic gender equity at a single point in time. However, any alternative measure – such as the mean proportion of husbands' housework over 2010 to 2015 – produces problems of endogeneity. We evaluate the robustness of our approach by repeating the analysis using data from 2011 on husband's proportion of housework, and data on fertility outcomes

from 2012 to 2015. Three models (Models (7)-(9)) are estimated based on the following equation:

$$ln\left(\frac{Pr(y_{i,2011-2015}=1|\boldsymbol{x}_{i,2010})}{1-Pr(y_{i,2011-2015}=1|\boldsymbol{x}_{i,2010})}\right)=\boldsymbol{x}'_{i,2010}\boldsymbol{\beta}. \tag{3.3}$$

In Equation 3.3, $y_{i,2011-2015}$ is a binary variable that equals 1 if a couple had one or more children from 2011 to 2015; $\boldsymbol{x}_{i,2010}$ is the vector of independent variables from 2010; and $\boldsymbol{\beta}$ is its associated vector of coefficients. In vector $\boldsymbol{\beta}$, the main coefficient of interest is husbands' proportion of housework. We expect our estimate for this coefficient to be positive, indicating that husbands' proportion of housework has a positive effect on realised fertility in the long-term. We also expect this estimate to be larger than the estimate for husbands' proportion of housework in Equation 3.1, which would indicate that husbands' proportion of housework has a greater effect on realised fertility in the long-term than in the short-term.

3.4.2 Variables

Housework balance is measured by the total hours of housework dedicated per week by each spouse. The PFSD questionnaires do not ask specific questions about different forms of domestic work, such as cleaning and childcare. The only questions about domestic labour concern the number of hours of housework dedicated per week by the respondent, and the number of hours dedicated per week by the respondent's spouse. Broadly, "housework" here includes activities such as cooking, cleaning, and laundry, but not childcare. We focus on the share of housework done by each spouse, because domestic gender equity concerns whether the balance of household labour is considered fair, rather than concerning the absolute amount of housework. Changes in individual perceptions of fairness tend to operate according to changes in personal circumstances (Perales et al. 2015). Our key explanatory variable is the husband's proportion of housework:

$$share=\frac{hhw}{hhw+whw},$$

where *share* gives the husband's proportion of housework, *hhw* gives the husband's weekly hours of housework, and *whw* gives the wife's weekly hours of housework. Husband's weekly hours of housework and wife's weekly hours of housework were also included as variables in our models, in order to control for absolute number of hours of housework done by each couple.

We included a range of control variables for variations in husbands' and wives' income, education, work, age, and parity (i.e. birth order). In terms of parity,

our analytical sample is restricted to couples at parity 1 or above. Since two-child families are normative, we include a binary control variable to control for families at parities of two and over. This is because families with one child are likely to face an additional incentive to have another child (because of the two-child norm), relative to families who already have two or more children. Table 3.1 lists and provides summary statistics for these variables. It indicates that husbands do about 29.7% of housework on average. Of the 2,176 person-years in the sample, there were 146 births in the four years 2011, 2012, 2013, and 2015. The proportion of couples that had one or more births across all years from 2011 to 2015 was 0.296, a total of $544 * 0.296 = 161$ couples.

Table 3.1: Summary statistics for regression variables.

Variable	**Mean**	**Standard deviation**
Birth in following year (binary)	.0671	.250
One or more births in 2011-2015 (binary)	.296	.457
Youngest child younger than five	.518	.500
Wife or husband desires another child (binary)	.186	.389
Tertiary education – wife (binary)	.474	.499
Tertiary education – husband (binary)	.468	.499
Proportion of housework done by husband	.297	.207
Wife's weekly hours of housework	16.6	13.4
Husband's weekly hours of housework	7.20	8.43
Wife's age in years	34.1	4.05
Wife works (binary)	.734	.442
Any parents live with couple (binary)	.483	.500
Log income	10.8	1.74
Parity two or more (binary)	.704	.457
Pays for a nanny (binary)	.150	.357
Year of the Dragon (binary)	.25	.433

Notes: For all variables except "One or more births in 2011-2015," the sample size is 2,176 person-years, covering 544 respondents over the survey years 2010, 2011, 2012, and 2014. For "One or more births in 2011-2015," the sample size is 544.

Source: PSFD 2018.

The 146 births from 2011 to 2015 are negatively distributed over time, with 56 births in 2011, 44 births in 2012, 34 births in 2013, and 12 births in 2015. This pattern is to be expected since all individuals in the sample have at least one child in 2010, and couples are not likely to have more than two children.

3.5 Results and discussion

In this section, we first discuss our model results on realised births, before focusing on our fertility desire models. Finally, we consider some policy implications of our results.

3.5.1 Realised births

Table 3.2 presents the estimates of the hybrid logistic regression models based on Equation 3.1. Table 3.3 displays the estimates of the cross-sectional model based on Equation 3.3. Only key variables and significant variables are displayed in Table 3.2 and Table 3.3. Tables displaying the full range of variables included in the models are reported in the Appendix: hybrid model estimates are displayed in Table A1, and cross-sectional estimates are displayed in Table A2.

3.5.1.1 Short-term effects

We firstly focus on Models (1)-(3) in Table 3.2. These models report the year-on-year effects of covariates on the probability of couples having a child in the following year. Unexpectedly, within or between coefficients for the proportion of housework done by husbands are statistically insignificant across all three models. The lack of statistical significance for within coefficients may indicate that if a couple starts sharing housework more equally at some point in time, any effect on the probability of another birth in the next year is modest. The lack of statistical significance for between estimates, on the other hand, may indicate that couples which share housework more equally (in a given year) are no more likely to have a birth in the next year than couples that share housework less equally. In terms of Hypothesis 1, our results indicate that greater housework balance has little or no impact on the probability of future births.

These findings are inconsistent with prior work indicating that greater housework balance has a positive impact on realised fertility (Cooke 2009; Dommermuth et al. 2017; Nagase and Brinton 2017). These conflicting results may be due to differences in contextual factors, but may also reflect the fact that prior empirical analysis has conflated the effect of within and between variations in housework balance. Once these effects are separated, they seem to exert little influence on fertility, at least in the short-term.

It may be the case that the sample size is too small to detect a short-term effect of husband's proportion of housework on fertility, meaning that there are too few birth events to detect an effect. However, we do have enough births in the long-term models to detect an effect (see next section), and there are no

Table 3.2: Selected hybrid panel data logistic regression estimates of births, by survey respondent group.

Type of estimate	Variable	(1) All Log-odds	Odds ratio	(2) Females Log-odds	Odds ratio	(3) Males Log-odds	Odds ratio
Within estimate	Proportion of housework done by husband	0.983	2.672	1.365	3.916	-0.160	0.852
		(1.199)		(1.806)		(1.970)	
	Wife's hours of housework	-0.0183	0.982	-0.0353	0.965	-0.0137	0.986
		(0.0256)		(0.0430)		(0.0390)	
	Husband's hours of housework	-0.0462	0.955	-0.0754	0.927	-0.00756	0.992
		(0.0333)		(0.0517)		(0.0523)	
	Wife's age	0.621***	1.861	0.558*	1.747	0.951***	2.588
		(0.153)		(0.239)		(0.247)	
	Parity 2 or more	-7.841***	0.000	-8.555***	0.000	-8.866***	0.000
		(0.707)		(1.165)		(1.116)	
	Desires more children	-0.509	0.601	0.227	1.255	-1.669*	0.188
		(0.511)		(0.855)		(0.763)	
Between estimate	Proportion of housework done by husband	0.414	1.513	0.170	1.185	0.557	1.745
		(1.523)		(2.424)		(2.491)	
	Wife's hours of housework	-0.00325	0.997	0.0269	1.027	-0.0526	0.948
		(0.0280)		(0.0427)		(0.0443)	
	Husband's hours of housework	-0.0582	0.943	-0.0487	0.952	-0.0703	0.934
		(0.0437)		(0.0695)		(0.0740)	
	Wife's age	-0.0540	0.947	-0.0829	0.920	-0.0618	0.932
		(0.0456)		(0.0973)		(0.0592)	
	Parity 2 or more	0.165	1.179	0.0736	1.076	0.406	1.501
		(0.436)		(0.712)		(0.646)	
	Desires more children	0.545	1.725	1.019	2.770	-0.387	0.679
		(0.786)		(1.261)		(1.159)	
	Constant	-3.308		-2.937		-6.995+	
		(2.199)		(3.853)		(3.987)	
	N	2176		1040		1136	

Notes: Within estimates are derived using variation within individuals over time, and between estimates use variation across different individuals. Estimates of the log-odds, standard errors (in brackets), and odds ratios are reported to three decimal places. The odds ratios are the exponents of the log-odds. $+p < .10$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

more births in the long-term models than there are in the short-term models; this suggests that we do not have too few births in the short-term models. That being said, we cannot be sure that we have enough births in the short-term models, because the number of observations is larger, i.e. 2,176 rather than 544. Since we cannot rule out the problem that there might be too few birth events, we merely conclude that our short-term models do not provide any clear evidence for or against gender equity theory.

In terms of control variables, the age of the wife and whether a couple has two or more children return statistically significant coefficients for the within estimates for all three Models (1)-(3). The mean wife's age is 34.1 (Table 3.1), reflecting the postponement of having children towards older childbearing ages. The low probability of a third birth reflects our expectation that most couples will have a maximum of two children, given the two-child family norm. By contrast, between estimates for these variables are statistically insignificant, which may reflect greater similarity in the structure of Taiwanese families in our sample, in relation to wife's age and parity.

Desiring more births is statistically significant in Model (3) for the within estimate, and is negatively associated with realised fertility, but both within and between estimates have insignificant p-values for Models (1) and (2). The Model

(3) results indicate that if a couple's desire to have a child increases, the likelihood of having a child decreases. This contradicts previous evidence that desiring a child is a strong predictor of fertility (Thomson et al. 1990). Since this result is found in the model that uses only male respondents, this could be explained in terms of husbands' gender ideology: since husbands with a traditional gender ideology tend to both desire more children and do less housework, the wives of husbands who desire more children might be less inclined to have additional children (Scanzoni 1976).

3.5.1.2 Long-term effects

Greater housework in the short-term seems to have no influence on realised fertility, but housework division may have a cumulative effect on the probability of a second birth over time. Thus, differences in housework balance may not lead to having a child in the short-term but may do in the long run. Models (7)-(9) in Table 3.3 report the estimates for the long-term effects of housework balance on fertility. The estimates reveal a positive and significant effect of husband's share of housework on the probability of couples having another child for female respondents, but not for male respondents or all respondents. Support for Hypothesis 1 is thus only found in the cross-sectional model containing only female respondents. This suggests that perceptions of how much each partner contributes to housework varies significantly by gender, with wives perceiving husbands to be doing less housework than husbands perceive themselves to be doing (Cerrato and Cifre 2018).

The husband's housework proportion coefficient in Model (8) is the largest across all three models in Table 3.3. The estimated effect of a husband doing a greater share of housework on the probability of realised fertility reflects the magnitude of its influence. The mean value for this variable is 29.7% (Table 3.1); were this to shift to 50% – i.e. if housework were shared equally – then the odds of having another child would increase by $exp(0.203 * 3.560) = 2.060$, a 20.6% increase. This suggests that increasing domestic gender equity could have a large effect on transitions to 2nd and higher-order births. This result broadly concurs with the existing evidence on the effect of household labour on fertility behaviour, although the magnitude of the effect here seems to be much larger than previously found (e.g. Dommermuth et al. 2017).

The results for wife's and husband's hours of housework are mixed, varying with the gender of the respondent. In Model (9) (male respondents), the results indicate that as husbands increase their hours of housework, and as wives decrease their hours of housework, the likelihood of having a child increases. These results reflect a situation in which the husband takes on a greater proportion

Table 3.3: Selected cross-sectional logistic regression estimates of births, by survey respondent group.

Variable	(7) All		(8) Females		(9) Males	
	Log-odds	**Odds ratio**	**Log-odds**	**Odds ratio**	**Log-odds**	**Odds ratio**
Proportion of housework done by husband	1.051	2.861	3.560**	35.163	-2.058	0.128
	(1.040)		(1.246)		(1.556)	
Wife's hours of housework	-0.0546*	0.947	-0.0452	0.956	-0.0784**	0.925
	(0.0228)		(0.0467)		(0.0268)	
Husband's hours of housework	-0.00846	0.992	-0.0601+	0.942	0.0664+	1.069
	(0.0266)		(0.0309)		(0.0398)	
Wife's age	-0.286***	0.751	-0.397***	0.672	-0.270***	0.763
	(0.0537)		(0.0971)		(0.0653)	
Parity 2 or more	-2.202***	0.111	-2.035***	0.131	-2.563***	0.077
	(0.294)		(0.421)		(0.464)	
Desires more children	0.627	1.871	0.252	1.287	0.979+	2.662
	(0.400)		(0.621)		(0.580)	
Gender	-0.0793	0.924				
	(0.263)					
Constant	7.991***		10.29**		8.810***	
	(1.887)		(3.225)		(2.471)	
N	544		260		284	

Notes: The dependent variable here is whether the couple went on to have one or more children in 2011-2015. Estimates of the log-odds, standard errors (in brackets), and odds ratios are reported to three decimal places. The odds ratios are the exponents of the log-odds. $+p < .10$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

of housework, providing supporting evidence to Hypothesis 1 (that husbands' housework proportion is a positive predictor of realised fertility) and Hypothesis 1 (that husbands' housework proportion affects realised fertility more strongly in the long-term than in the short-term). However, the coefficient for husbands' proportion of housework is not statistically significant. Since the proportion of husbands' housework is a function of wives' hours of housework and husbands' hours of housework, if any two of these variables change then so must the third. This makes disentangling the effects of these three variables more complicated, although problems of multicollinearity are avoided by the fact that husbands' housework proportion is a nonlinear function of wives' hours of housework and husbands' hours of housework. In contrast to Model (9), Model (8) (female respondents) returns a large and positive estimate for husbands' proportion of housework, but then a significant and negative result for husbands' hours of housework, and no significant result for wives' hours of housework. The Model (8) results also imply that increasing husbands' hours of housework decrease the likelihood of having another child, if the proportion of husbands' housework remains the same (i.e. if wives' hours of housework increase proportionally).

Despite the strong effect found in Model (8), when the analysis was repeated for data in 2011 (predicting one or more births from 2012 to 2015), no such effect was found. This suggests that analysing the effect of domestic gender equality on fertility over longer periods is likelier to find a significant result, and that such effects may not be evident over time periods shorter than five years. It could

also be the case that, as for the short-term models, there was an insufficient proportion of couples having children at all over 2012 to 2015: while 29.7% of couples had at least one additional child between 2011 and 2015, only 23.2% of couples had at least one additional child between 2012 and 2015. In this way, the failure to replicate the 2010 result with data from 2011 could be due to the shorter time period, or could be due to an insufficient proportion of the sample having children to ensure enough variability to identify an effect, or could be due to a combination of these two factors.

In terms of control variables, the wife's age and the couple being at a parity of two or more were significant and negative across Models (7)-(9). The negative coefficient for wife's age reflects a high median wife's age of 32 in 2010, with half of the women in the sample being 37 by 2015. As 37 is towards the end of most women's childbearing years, these older women would have likely had less children over 2011-2015 than women who were younger in 2010, reflecting wife's age as a negative predictor of fertility in our models. The coefficients for parity of two or more are similar to those for the short-term within estimates, although their magnitude is smaller. This difference could be due to the fact that in the short-term models, parity 1 couples who had more children between 2011 and 2015 would subsequently have been counted as a parity 2 or more couple. In the long-term models, these same couples were only counted as parity 1 couples, as covariate data from only 2010 was used. Overall, these control variable coefficients suggest that couples with older wives and with two or more children in 2010 were less likely to have another child across 2011-2015. The desire for more children is a positive predictor of the likelihood of having another child in all three models, indicating that couples who desire more children go on to have more children, as expected.

Our results suggest that domestic labour balance effects fertility in the long-term, but not in the short-term. We have a positive and significant estimate for the proportion of husbands' housework in Model (8), but no such estimate in the short-term models. This supports our earlier argument that perceptions of gender equity in household roles may become less responsive to changes in realised housework division over time. Even though a husband might actually have been doing more housework in recent years, a wife might still regard her expected role as being fundamentally the same as that established during their early relationship and the early years of their first child. Moreover, those beliefs could be largely unconscious – a general feeling about how things are done in that particular family, rather than a conscious belief about how domestic labour would be shared after a potential future birth.

These considerations suggest that future research in this area should inves-

tigate the relationship between housework and fertility over different periods of time. We have couched our argument here in terms of marital satisfaction – it still needs to be shown whether domestic labour equity effects fertility in a similar manner over time. Specifically, research is needed on how equity effects fertility at different times in a couple's life, and on how equity at a given point in time affects fertility over the short-term and the long-term.

3.5.2 Desire for more births

Table 3.4 presents the estimates of the hybrid logistic regression models based on Equation 3.2. As in the previous subsection, only a selection of all the explanatory variables are displayed; the full range of variables included in the models are reported in the Appendix, in Table A1.

The results reveal no significant effect for any of the six housework proportion variables. These results suggest that more balanced housework arrangements do not necessarily lead to a greater desire for having a child, providing evidence against Hypothesis 2 (that husbands' share of housework is positively associated with the desire for more children). However, the within coefficients reveal a significant positive effect for husbands' hours of housework, when all respondents are used (Model (4)) and when only female respondents are used (Model (6)). This indicates that wives express a greater desire to have children as their husbands start doing more housework. Moreover, our results for the effect of hours of husbands' housework are consistent with cross-sectional analyses for Taiwan, which find that husbands' increased share of housework positively impacts couples' desire for another child (Kan and Hertog 2017). We therefore find partial support for gender equity theory, since husbands doing more hours of housework is associated with an increase in wives desiring more children.

In terms of control variables, there are significant negative coefficients for wives' age, for both within- and between-estimates. These results suggest that younger wives have a greater desire for more children than older wives, and that wives' desire for children decrease as they age. Since the wives in our sample have a mean age of 34.1 (Table 3.1), which is towards the older end of the child-bearing ages, the negative effects of ageing on fertility desire are to be expected, as wives start to exit their child-bearing period. Moreover, as wives age they will be more likely to have the children they desire, and then not desire children thereafter. Within estimates for parity 2 or more are not statistically significant in any model, indicating that couples progressing from parity 1 to parity 2 do not desire more children. Additionally, the between estimates for parity 2 or more are statistically significant and negative, indicating that couples with 2 or more

Table 3.4: Selected hybrid panel data logistic regression estimates of desire for more births, by survey respondent group.

Type of estimate	Variable	(4) All Log-odds	Odds ratio	(5) Females Log-odds	Odds ratio	(6) Males Log-odds	Odds ratio
Within estimate	Proportion of housework done by husband	-0.685	0.504	-1.281	0.278	0.352	1.422
		(0.833)		(1.159)		(1.264)	
	Wife's hours of housework	0.00560	1.006	0.0163	1.016	0.00116	1.001
		(0.0146)		(0.0230)		(0.0196)	
	Husband's hours of housework	0.0307	1.031	0.0533+	1.055	-0.0000332	1.000
		(0.0206)		(0.0279)		(0.0319)	
	Wife's age	-0.616***	0.540	-0.683***	0.505	-0.576***	0.562
		(0.0949)		(0.142)		(0.137)	
	Parity 2 or more	-0.502	0.605	-0.372	0.689	-0.793+	0.452
		(0.327)		(0.492)		(0.454)	
Between estimate	Proportion of housework done by husband	-0.110	0.896	1.502	4.491	-2.694	0.068
		(1.377)		(1.740)		(2.290)	
	Wife's hours of housework	-0.00138	0.999	0.0289	1.029	-0.0137	0.986
		(0.0234)		(0.0344)		(0.0332)	
	Husband's hours of housework	0.0139	1.014	-0.0165	0.984	0.0909	1.095
		(0.0389)		(0.0506)		(0.0645)	
	Wife's age	-0.135***	0.874	-0.219**	0.803	-0.119*	0.888
		(0.0403)		(0.0746)		(0.0500)	
	Parity 2 or more	-3.012***	0.049	-2.874***	0.056	-3.391***	0.034
		(0.323)		(0.437)		(0.497)	
	Constant	1.141		4.250		-3.044	
		(1.846)		(2.827)		(3.063)	
	N	2176		1040		1136	

Notes: Within estimates are derived using variation within individuals over time, and between estimates use variation across different individuals. Estimates of the log-odds, standard errors (in brackets), and odds ratios are reported to three decimal places. The odds ratios are the exponents of the log-odds. $+p < .10$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

children are less likely to want more children than couples with 1 child. This is expected, given the norm of 2 child families.

3.5.3 Policy implications

In policy terms, these results suggest that encouraging equalisation of household labour could play a role in increasing fertility. The Taiwanese government first introduced pro-natalist policies in 2006 and 2008, which focused on financial support and childcare for couples with young children (Chen 2012; Frejka et al. 2010; Lee 2009; Lee and Lin 2016). Parental leave of up to two years was introduced slightly earlier, but only became paid (60% of the mother's salary during the first six months) from 2009 (Lee and Lin 2016; Tsai 2012). Fathers are entitled to three days of leave on full pay, and can share in the two years of leave, again paid at 60% of their salary during the first six months (Tsai 2012).[3] Promoting domestic gender equity has not been a part of any pro-natalist policy so far – for a recent overview, see Yip and Chen (2016).

In Scandinavian countries, paternity leave has long been regarded as a means of promoting domestic gender equality (Cools et al. 2015; Kotsadam and Finseraas 2011). These countries are also characterised by stable and near-replacement level fertility; however, causal links between policies and fertility are potentially

[3]https://www.bli.gov.tw/en/0013195.html

confounded by gender equitable attitudes across these societies at large. Nonetheless, there is some evidence – both in Scandinavia and in other developed countries – that the availability and uptake of paternity leave has a positive effect both on subsequent domestic gender equality, and on second births (Boll et al. 2014; Kotsadam and Finseraas 2011; Lappegård 2010). For Taiwan, long periods of paid paternity leave could potentially be a viable means of promoting fertility. Providing longer periods of better-paid paternity leave could encourage fathers to perform a larger share of housework and childcare, leading to more gender-equitable outcomes at home and subsequent increases in fertility.

3.6 Conclusion

Taiwan has the third lowest TFR in the world and has remained persistently low over the last two decades. Identifying the key factors underpinning these patterns is important to develop effective ways to promote fertility. In this paper we sought to establish whether the division of household labour exerts a significant influence on shaping the probability of desiring another child, or on having another child. Drawing on panel data from the PSFD, we examined the short- and long-term effects of household work on fertility behaviour.

The evidence accumulated here shows no evidence of a short-term effect of housework division on fertility behaviour or fertility desire. A greater share of household labour taken up by husbands is not statistically associated with having an additional child or wanting to have another child. Yet, men doing a greater share of domestic labour results in a higher probability of having an additional child in the long-term. Specifically, couples in which husbands consistently do a greater share of housework are more likely to have an additional child in the next five years than couples in which wives take a larger responsibility for domestic work. Our results indicated that if a husband increases his share of housework from a sample mean of 0.297 to 0.5, the odds of having another child increase by 20.6% over a five year period. This finding is indifferent to the absolute number of hours dedicated to housework by either partner.

These results comprise the first empirical support for the validity of gender equity theory as an explanation for fertility behaviour in Taiwan. Our conclusions are consistent with existing evidence in Europe and North America, and elsewhere in East Asia (Balbo et al. 2013; Kan and Hertog 2017; Oláh 2003). Yet, the magnitude of the long-term effect found in this study is larger. Prior work tends to focus on examining the short-term effects of housework on fertility behaviour. Further research is needed to establish whether domestic gender equity acts more strongly on fertility in the long-term than in the short-term, and whether the

effects of domestic gender equity are greater in Taiwan – and more widely across East Asia – than in Europe and North America.

Our findings suggest that governments seeking to raise fertility should focus on policies aiming at promoting greater domestic gender equality, such as increased paternity leave and child care provision. Existing evidence points to the effective use of paternity leave policies seeking to increase fertility in Scandinavian countries. However, research is needed to assess if recent increases in parental and paternity leave in East Asia are having such an effect.

Our findings also suggest policies aimed at promoting fertility should remain in place for some years as the effects of greater housework balance are not immediate. They may take up to five years to be realised. At the same time, policies should also be targeted at teaching gender equality. Children could be taught at school about the importance of sharing all forms of labour equally between the sexes. Such policies are standard in many near-replacement fertility developed countries, such as in Sweden, where pre-school gender equality teaching projects have been implemented since the 1990s (Bayne 2009).

4

Paper 2 – The Effect of Parental Leave Policies on Increasing Fertility: A Systematic Review

Low fertility is set to worsen economic problems in many developed countries, and parental leave has emerged as a key pro-natal policy. However, the literature contends that evidence for the effect of parental leave on fertility is mixed. We conduct the first systematic review on this topic. By applying a rigorous search protocol, we identify and review empirical studies that quantify the impact of parental leave policies on fertility. We focus on experimental or quasi-experimental studies that can identify causal effects. We identify 11 papers published between 2009 and 2019, evaluating 23 policy changes across Europe and North America from 1977 to 2009. Results are a mixture of positive, negative, and null impacts on fertility. To explain these apparent inconsistencies, we propose a new conceptual framework which decomposes the total effect of parental leave on fertility into the "current-child" and "future-child" effects. We decompose these into effects on women at different birth orders, and specify types of study design to identify each effect. We classify the 23 studies in terms of the type of effect identified, revealing that all the negative or null studies identify the current-child effect, and all the positive studies identify the future-child or total effect. Since the future-child and total effects are more important for promoting aggregate fertility, our findings show that parental leave does in fact increase fertility when benefit increases are generous. Furthermore, our conceptual framework provides a new way of understanding and classifying the effects of pro-natal policies on fertility. Additionally, we propose ways to adapt the ROBINS-I tool for evaluating risk of bias in pro-natal policy studies.

4.1 Introduction

Persistent low fertility causes economic problems and is associated with parents having fewer children than they want. Many countries in East Asia and Europe have fertility levels significantly below the replacement level of 2.1 births per woman, and countries across Asia and Latin America are projected to join them in the coming decades (UN 2019). When low fertility rates persist, rising old-age dependency ratios and reductions in the working-age population cause higher consumption, lower investment, and economic stagnation or decline (Bloom et al. 2010; Caldwell et al. 2002; McDonald 2008). On the individual level, low fertility can be indicative of people having fewer children than they would ideally like to have (Beaujouan and Berghammer 2019; Chen and Yip 2017; Spéder and Kapitány 2014). The failure of individuals to fulfil their childbearing desires can negatively impact emotional well-being (Casterline and Han 2017; Priebe 2020; Ugur 2020).

Governments try to promote fertility through family policies (Gauthier 2007; Raute 2019; Rindfuss and Choe 2016). Family policies aim to increase fertility either through helping parents balance work and family, or through reducing the costs of childbearing and childrearing (Gauthier 2007, 2008; Rindfuss and Choe 2016). Family policies can be categorised as child-related cash transfers, childcare subsidies, or financial support through the tax system (OECD 2019c). Child-related cash transfers tend to take up the largest proportion of public expenditure, amounting to 1.3% of GDP across OECD countries in 2015, a total spend of over $0.75 trillion (OECD 2019a,b,c). Within child-related cash transfers, parental, maternity and paternity leave policies (henceforth 'leave') refer to state mandated arrangements for parents to take time off work during pregnancy or after childbirth. In 2015, average OECD public expenditure on leave was $12,100 per infant (at purchasing power parity, 2010 USD) (OECD 2019d). Understanding the extent of impacts of such policies on fertility is thus critical to guide pro-natal policy efforts.

Low fertility is particularly acute in East Asia, in the advanced economies of South Korea, Taiwan, Japan, and Singapore (UN 2019). In these societies, which share a common Confucian heritage and similar development trajectories, governments have aimed to increase fertility through a range of pro-natal policy efforts (Suzuki 2013). These countries differ from low fertility countries in the Europe, North America and Oceania in the set-up of their welfare systems and orientations towards government intervention in the society and economy (Esping-Andersen and Billari 2015). Therefore in order to evaluate the efficacy of pro-natal policies in general, consideration of the experience of East Asian countries must

be taken into account.

This paper focusses on parental leave – rather than other family policies – because leave can uniquely increase domestic gender equity, and because domestic gender inequity has been identified as a key cause of low fertility (Goldscheider et al. 2015; McDonald 2006; Tamm 2019). Over the past two decades, researchers have increasingly argued that the tension between increasing career ambitions and persistently gendered domestic obligations has meant that women's ability to reconcile work with family life has become more restricted, which in turn seems to have contributed to low fertility (e.g. Baizan et al. 2016; Duvander et al. 2019; Meier and Rainer 2017). Leave can help equalise the division of domestic and formal labour between men and women through enabling mother' s return to work, and through encouraging fathers to do more housework and childcare [1] (Baum and Ruhm 2016; Pronzato 2009; Tamm 2019). Through equalising the gender balance of labour, leave can reduce the cost of childrearing for women, and thereby increase fertility (Baizan et al. 2016; Kotsadam and Finseraas 2011). Although gender equity is our motivation for focussing on leave, we only seek to establish whether leave has an impact on fertility.

Whether leave actually increases fertility remains a matter of debate (Balbo et al. 2013; Hoem 2008; Olivetti and Petrongolo 2017). Gauthier's (2007) review of the effect of family policies on fertility found the evidence to be "mixed," an evaluation echoed across the empirical literature (e.g. Hong and Sullivan 2016; Lappegård 2010; Matysiak and Szalma 2014). Similarly, the review by Bergsvik et al. of the effect of family policies on fertility (2020) concludes that the effect of leave on fertility is ambiguous. There are certainly cases in which governments have provided generous leave policies and low fertility has persisted, such as in Slovenia (Stropnik and Šircelj 2008). There are also cases in which generous new leave policies have been accompanied by large increases in fertility (e.g. East Germany), and cases in which generous leave policies have been accompanied by stable and high fertility (e.g. Czechoslovakia and Sweden) (Buttner and Lutz 1990; Hoem 1990, 1993, 2005; Monnier 1990; Salles 2006). However, most empirical studies use methods that prohibit identification of a causal effect of leave on fertility. To date, there have been no peer-reviewed systematic reviews focussing on leave and fertility which discriminate between studies that can identify causal effects and studies that cannot. Whether leave does in fact cause higher fertility therefore remains an open question. A peer-reviewed systematic review could reliably answer this question by showing whether or not more generous parental leave leads to higher fertility, thereby resolving existing academic debates and

[1] In some countries such as Sweden, gender equality is an explicit goal of leave policy (Duvander and Johansson 2012).

giving governments a sounder footing for policy-making.

This paper provides the first peer-reviewed systematic review of the effect of leave on fertility. We aim to find, evaluate, and synthesise all relevant, experimental or quasi-experimental studies, in a rigorous, transparent, and reproducible fashion. We seek to answer the question: to what extent does leave increase fertility? Since we are interested in informing pro-natal policies in low-fertility countries, we restrict our search to countries and time periods where fertility is persistently below 2 births per woman (broadly speaking, in high-income countries from the 1950s onwards). Using a thorough search of all published English-language material catalogued online, we identify 11 papers that match our inclusion criteria, containing 23 different studies that can plausibly test for a causal effect of leave on fertility.[2]

Our paper makes three contributions. Firstly, we develop a new conceptual framework (Section 4.4.2), building on Lalive and Zweimüller's (2009) definitions of the "current-child effect" and the "future-child effect". Our framework decomposes the impact of leave on individuals by type of effect and by parity, specifies study designs that can be used to identify each impact, and can be applied to assess the impact of other pro-natal policies on fertility. Secondly, we find that leave reforms which provide generous increases in duration or remuneration consistently increase fertility, implying that large increases in leave benefits are a viable strategy for governments seeking to raise fertility (Section 4.4). Categorising the 23 studies in terms of our framework reveals that studies whose methods identify a broader class of effects consistently find positive results, and that all negative or null studies only address a narrow class of effects that are of marginal interest to policy-makers. Thirdly, we propose ways to adapt the ROBINS-I tool for the evaluation of studies of pro-natal policies (Section 4.4.3). ROBINS-I was designed as a tool for assessing risk of bias (RoB) in non-randomised studies of medical interventions (Sterne et al. 2016); we identify three key reasons why it is not directly applicable to studies of public policy interventions. First, ROBINS-I assumes the existence of placebo effects, which do not exist in a public policy context. Second, the notion of an "intention-to-treat" (ITT) study is problematic in the case of public policy because individuals can self-select into being (in)eligible for a policy after their initial assignment (in medical studies, participants cannot change their assignment status after being assigned or not assigned to treatment). Third, ROBINS-I does not distinguish between policy eligibility and policy availability, which comprise critical dimensions of leave policies. As well as identifying

[2]By "different studies" we mean individual models (or sets of related models) that are different either in terms of the specific policy being evaluated, or different in terms of the analytical sample (e.g. a paper that has separate models for the effect of a policy on women at parity 1 and parity 2 counts as having two different studies).

these three considerations and proposing strategies to account for them, we also identify the sources of bias inherent in each of the study designs in our conceptual framework, providing future empirical researchers with a checklist of sources of bias and strategies to minimise RoB.

The rest of this paper is structured as follows. Section 4.2 provides background information on the economic mechanisms by which leave is theorised to affect fertility, and introduces Lalive and Zweimüller's (2009) definitions of the current-child and future-child effects. Section 4.3 describes systematic review methods and explains how we conducted this review. Section 4.4 presents the results of our search and filtering process, and develops our conceptual framework to classify and assess studies and the type of fertility effect they capture. Section 4.4 also assesses the RoB of included studies, and synthesises the study findings. Section 4.5 discusses our analysis and findings, and Section 4.6 offers some final concluding remarks.

4.2 Background

This section provides background information on two topics. Section 4.2.1 describes the economic channels by which leave is theorised to increase fertility, and Section 4.2.2 provides an exposition of Lalive and Zweimüller's (2009) theory of the current-child effect and the future-child effect.

4.2.1 Economic theory of how leave increases fertility

Leave increases fertility by enabling parents to balance work and childcare, and by lowering the net costs of childrearing (Gauthier 2008). In the absence of leave, a parent broadly has two options after the birth of their child:

(1) Return to work soon after childbirth. Lose little to no job income, but spend money on childcare after returning to work.

(2) Quit work in order to look after the child. Lose all job income, but save money on childcare.

Leave presents a third option:

(3) Take some partially remunerated weeks or months off, and then return to work. Lose some job income, but spend no money on childcare for the duration of leave.

By enabling parents to both take care of their child and return to their job, leave represents a middle-ground between options (1) and (2). Prospective parents may want to spend time with their new-born infant and save money on childcare (prohibiting option (1)), but also want to keep their job and avoid negative impacts on their career (prohibiting option (2)) (Nagase and Brinton 2017; Wood et al. 2016). Prospective parents with these preferences may be less likely to have a child in the absence of leave, and so the availability of leave may cause them to have a child they might not have had otherwise (Becker 1973; Ermisch 2003). Regardless of parental preferences, option (3) might have the lowest net financial cost. If the rate of remuneration of leave is high and childcare is expensive, then taking leave will be cheaper than either other option (Gauthier 2008). Leave can therefore make childrearing less costly, facilitating childbearing.

4.2.2 Effects on fertility: current-child effect and future-child effect

Leave can affect fertility behaviour in two key ways. In a seminal analysis in the econometrics literature, Lalive and Zweimüller (2009) distinguish between the "current-child effect" and the "future-child effect" of a leave policy on fertility. This conceptual framework has become commonly used in other studies, in order to elucidate which type of effect is being identified (e.g. Cygan-Rehm 2016; Dahl et al. 2013; Raute 2019).

The current-child effect refers to the effect of being able to take more leave for the child just born, on subsequent fertility. Women giving birth shortly before and shortly after a reform receive different benefits for the child they just had. However, both groups of women will receive equal benefits for any subsequent children. Therefore if there is any long-term difference in fertility between the two groups, it must be due to the different benefits they received for the child born around the time of the reform.

The future-child effect refers to the effect of a greater amount of leave available in the future. It is called the "future" child effect to distinguish it from the current-child effect, and captures the idea that, if a woman knows she will receive more generous leave entitlements if she has a child, she will probably be likelier to have that child. Lalive and Zweimüller identify the future-child effect by comparing the fertility of mothers in the years before the reform, with the fertility of mothers in the years after the reform.

Lalive and Zweimüller go on to argue that the sum of the current-child effect and the future-child effect gives the total effect of leave on fertility. However, we note that their study design can only identify effects on women at parities of

1 or higher, since the women in their samples all had at least one birth. This means that their study design cannot identify the effect of leave on women with no children, and therefore cannot identify the total effect of leave on fertility.[3] While other studies have used Lalive and Zweimüller's terminology, no study has highlighted the role of parity in classifying effects. Separating the impact of a leave policy on women at different parities can enable us to understand the different processes by which individuals choose to build a family.

4.3 Methods

Our methods were guided by the policies and guidelines of The Campbell Collaboration for conducting systematic reviews (https://campbellcollaboration.org/) (Campbell Collaboration 2019). Prior to conducting literature searches we produced a review protocol, which was registered and published online at the International prospective register of systematic reviews (PROSPERO), on the 9th of September 2019 (Thomas et al. 2019). Full details of our method are given in the review protocol, a version of which is given in Appendix B.1. This report was written according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist (Moher et al. 2009). All deviations from the protocol are specified in Appendix B.2.

4.3.1 Criteria for inclusion of studies

Studies must be primary, empirical, quantitative studies that assess the effect of a leave policy on fertility at the micro-level. We exclude macro-level studies since they cannot identify the causal mechanisms by which policy affects fertility (Neyer and Andersson 2008). We are interested in the effects of policy changes involving one or more of maternity, paternity, and parental leave. Changes can either be in increasing, decreasing, or restructuring leave. Since the state is the policy-maker of interest, we are not interested in the effects of firm-specific policies on fertility. Furthermore, since we aim to collect evidence to inform policy in low-fertility countries, we are only interested in finding studies of policies implemented in countries with a TFR below 2. We apply this restriction because we want the included studies to have external validity in terms of taking place in settings that are comparable to other countries with low fertility (Shadish et al. 2002). Countries and time periods that are eligible for inclusion are specified in the protocol. In terms of fertility, we are interested primarily in quantum effects

[3]Alternatively, Lalive and Zweimüller's total effect (for women at parities of one or more) can be regarded as an underestimate, or lower bound, on the total effect (for women at parities of zero or more).

rather than tempo effects. We exclude articles that purely consider policy effects on birth timing or seasonality.

We only include studies with strong designs that can plausibly provide evidence of a causal relationship between leave changes and fertility. These study designs are: randomised control trials (RCTs), quasi-experimental designs, and natural experiments, which can all provide evidence of a causal relationship through specifying appropriate means of estimating counterfactual situations. These designs therefore have a higher degree of internal validity than purely correlational or observational studies (Shadish et al. 2002). A key element in identifying a cause is establishing that the cause occurred before the effect, and so we are interested only in the effects of leave policy changes on fertility, rather than any relationship between existing leave policies and fertility.

Finally, we only consider studies that evaluate the effect of leave availability on fertility, rather than studies that evaluate the effects of leave uptake or use on fertility. This decision is motivated by the Cochrane Collaboration's advice that evaluating "intention-to-treat" (ITT) effects tends to result in less biased outcomes than "per-protocol" effects (Higgins and Green 2008). Since parents self-select into using leave, parents who use leave are likely to differ systematically from those who do not, meaning that the causal effect of leave on fertility cannot be identified (Lappegård 2010). Moreover, the availability of leave can motivate parents to conceive even if they do not take leave after childbirth. This means that only examining leave use cannot capture the full effect of the policy on fertility, which is of most interest to policy-makers looking to increase fertility.

4.3.2 Search strategy and filtering process

A flowchart illustrating our search and filtering process is given in Figure 4.1. The search process was divided into searching academic databases, searching grey literature sources, hand-searching relevant journals, and snowball searching using the references and citations of included articles. Filtering of included articles was done on the basis of titles, then abstracts, then full texts. Appendix B.3 details the search procedure for one database, "Academic Search Complete."

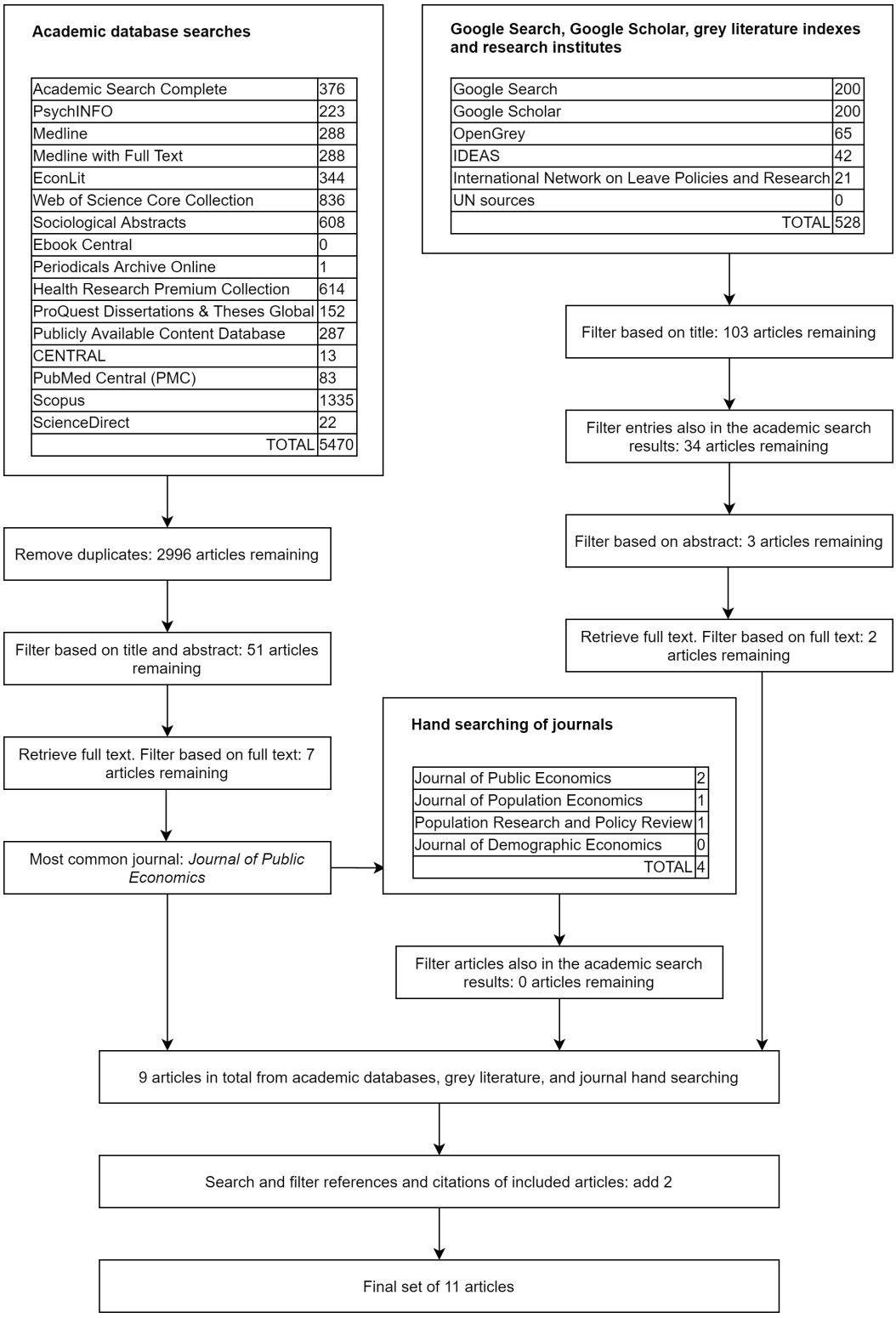


Figure 4.1: The literature searching and filtering process.

4.3.3 Assessing study quality

To assess study quality we used the ROBINS-I tool for assessing Risk of Bias (RoB) in non-randomised studies (Sterne et al. 2016). ROBINS-I is applied separately to each study in the review, and works by comparing the (non-randomised) study to a hypothetical, idealised RCT, in which there would be no RoB. Using ROBINS-I involves answering a series of signalling questions on characteristics of each study, across seven domains of bias: confounding, selection bias, misclassification bias, performance bias, bias from missing data, detection bias, and outcome reporting bias (these domains are explained more fully in Appendix B.4). The answers to the signalling questions are used to generate an overall classification of the study's RoB, ranging from "low risk of bias" (where the study is considered to be comparable to a well-performed RCT), to "critical risk of bias" (where the study is too problematic to provide any useful evidence, and should not be included in the synthesis). The signalling questions for each domain are determined by whether the study aims to measure the effect of "assignment to intervention," or the effect of "assignment and adherence to intervention."

4.3.4 Method of synthesis

Due to the many dimensions of difference between leave policies, it would be inappropriate to attempt any kind of statistical meta-analysis of the study results. Instead, we conducted a narrative review of included studies, using the Economic and Social Research Council's (ESRC) "Guidance on the Conduct of Narrative Synthesis in Systematic Reviews" document (Popay et al. 2006).

4.4 Results

This section provides an overview of the 11 papers and 23 studies from our search, and shows their results to be mixed. But, we reveal that there is a key underlying reason for the apparently mixed evidence. To this end, we develop a new conceptual framework for understanding the different fertility effects identified by the studies. We use this framework to classify the 23 studies, and then evaluate their RoB. Finally we synthesise the results.

4.4.1 Overview

The search and filtering process – with the number of articles removed at each stage – is summarised in Figure 4.1. We separately searched for articles in both

the academic and grey literature. We first scanned the academic literature, returning 5,470 results; after removing duplicates, 2,996 remained. Filtering based on title and abstract reduced these to 51 papers. After careful reading of the full text, this sample was further reduced to 7 papers. Secondly, we swept the grey literature returning 528 results. Filtering based on title and abstract reduced these to 3 results, and after reading the full texts 2 articles were preserved for the final analysis. We also conduced hand searches of relevant journals, but this returned no new results. We then conducted a snowball search of references and citations, which returned 2 further results. Thus 11 was the final number of articles included for analysis. The articles were: Ang 2015; Cannonier 2014; Carneiro et al. 2015; Cools et al. 2015; Cygan-Rehm 2016; Dahl et al. 2013; Duvander et al. 2016; Farré and González 2019; Hart et al. 2019; Lalive and Zweimüller 2009; Raute 2019.

Table 4.1 summarises key information on the 11 papers and 23 studies, including author names and publication year, year of the reform, country, type of leave, sample size, dependent variable, and effect found. The focus of the studies varies slightly. Two studies evaluated maternity leave; one study evaluated paternity leave; and 19 studies evaluated parental leave –although these policy changes differed in whether they affected mothers or fathers in practice. One study evaluates a reform that included all three of parental, paternity and maternity leave. The studies cover 7 countries across Western Europe and North America. The vast majority of studies evaluate reforms implemented in the 1990s and 2000s, with only three studies evaluating reforms before 1990. The length of follow-up varies widely, from 2 years to 30 years.

Eight studies found a positive effect of leave on fertility, one found a negative effect, and 14 found no evidence of an effect. It therefore would seem like the evidence on the effect of leave on fertility is mixed. However, Table 4.1 does not indicate which type of effect –current-child or future-child –each study identified. The reason for this is that some of the included studies measure effects that cannot be identified in Lalive and Zweimüller's original classification.

Table 4.1: Key characteristics of included studies.

Author(s) (year of publication)	**Year of reform**	**Country**	**Type of leave**	**Sample size**	**Dependent variable**	**Effect of the new policy (interpretation)**
Ang (2015)	2006	Canada	Maternity, paternity and parental	2,048,800	Child born in Quebec up to 2008	0.012^{***} (women 23.53% likelier to have a birth)
Cannonier (2014)	1993	USA	Maternity	11,963	First child born in the next 17 years	0.0519^{**} (women 5.19 percentage points likelier to have a birth)
				22,263	Second child born in the next 17 years	0.0296^{*} (women 2.96 percentage points likelier to have a birth)
Carneiro, Løken, & Salvanes (2015)	1977	Norway	Parental (maternity in practice)	63,571	Number of children born in next 30 years	-0.011 (women had 0.011 less children)
Cools, Fiva, & Kirkebøen (2015)	1993	Norway	Parental (paternity in practice)	23,985	Child born in the next 14 years	0.004 (fathers had 0.004 more children)
Cygan-Rehm (2016)	2007	Germany	Parental	39,826	Child born in next 5 years	0.046 (women 4.6 percentage points likelier to have a birth)
Dahl et al. (2016)	1987	Norway	Parental (maternity in practice)	17,580	Number of children born in the 14 years after each reform	-0.030 (women had 0.030 fewer children)
	1988			19,310		0.017 (women had 0.017 more children)
	1989			20,093		-0.016 (women had 0.016 fewer children)

Table 4.1 cont.

Author(s) (year of publication)	**Year of reform**	**Country**	**Type of leave**	**Sample size**	**Dependent variable**	**Effect of the new policy (interpretation)**
Dahl et al. (2016)	1990	Norway	Parental (maternity in practice)	21,508		−0.026 (women had 0.026 fewer children)
	1991			21,717		−0.024 (women had 0.024 fewer children)
	1992			21,838		0.042* (women had 0.042 more children)
Duvander, Johansson, & Lappegard (2016)	1993	Norway	Parental (paternity in practice)	1,804	Second child born in the next 10 years	0.98 (couples 2 percentage points less likely to have a birth)
				1,702	Third child born in the next 10 years	0.94 (couples 6 percentage points less likely to have a birth)
	1995	Sweden	Parental (paternity in practice)	3,880	Second child born in the next 10 years	0.95 (couples 5 percentage points less likely to have a birth)
				3,383	Third child born in the next 10 years	0.94 (couples 6 percentage points less likely to have a birth)
Farré & González (2019)	2007	Spain	Paternity	(not reported)	Child born in the next 6 years	−0.0179*, though not robust to bandwidth (fathers 5.3% less likely to have a birth)

Table 4.1 cont.

Author(s) (year of publication)	**Year of reform**	**Country**	**Type of leave**	**Sample size**	**Dependent variable**	**Effect of the new policy (interpretation)**
Hart, Andersen & Drange (2019)	2009	Norway	Parental (paternity in practice)	9,757	Child born in the next 5 years	−0.04 (couples had 0.04 fewer children)
Lalive & Zweimüller (2009)	1990	Austria	Parental (maternity in practice)	6,180	Child born in the next 10 years	0.035** (women 3.5 percentage points likelier to have a birth)
				5,997	Child born in the next 3 years	0.068** (women 6.8 percentage points likelier to have a birth)
	1996		Parental (maternity in practice)	(not reported)	Child born in the next 3 years	−0.001 (women 0.1 percentage points less likely to have a birth)
				(not reported)	Child born in the next 3 years	Positive effect (numerical result not reported)
Raute (2019)	2007	Germany	Parental	644,981	Child born to affected women after the reform	6.424*** (women 16% likelier to have a birth)

Note: For the results in the final column, * indicates $p < 0.1$, ** indicates $p < 0.05$, and *** indicates $p < 0.01$.

4.4.2 Effects of leave on fertility: conceptualisation and identification

We now develop a formal conceptual framework to assess the effect of leave on fertility, building on Lalive and Zweimüller (2009). Section 4.4.2.1 extends their terminology to include effects on individuals at all parities, Section 4.4.2.2 explains the empirical strategies used to identify each effect, and Section 4.4.2.3 classifies the 23 studies in terms of our framework.

4.4.2.1 Types of effects of leave on fertility

As discussed in Section 4.2, Lalive and Zweimüller argue that the total effect of the leave policy on fertility is the sum of the current-child effect and the future-child effect. However, since their analytical sample only consisted of women who have already had at least one child, they could not identify the effect of the policy on women at parity 0. For the purposes of this review, we use the term "future-child effect (parity 1+)" in the sense that Lalive and Zweimüller use "future-child effect," to mean an effect on women at parities of 1 or higher. In order to explain studies that evaluate the effect on women at parity 0, we introduce the term "future-child effect (parity 0)" to mean the effect of the policy on women who have not had any children. We use the term "total effect (parity 1+)" to mean the total effect identified by Lalive and Zweimüller, and "total effect" to mean the sum of all these effects across the population, as displayed in Figure 4.2 below.

We contend that pro-natal policy-makers are more interested in the future-child effect and the total effect, than they are in the current-child effect. In the language of experimental design, the future-child effect and the total effect have high "construct validity," and the current-child effect has low construct validity (Shadish et al. 2002). "Construct validity" refers to whether the specific features of an experiment validly capture the underlying concepts, or "target constructs" (ibid.). When designing a leave policy for pro-natal purposes, policy-makers have some concept of "the effect of leave on fertility." We contend that policy-makers have one of two conceptualisations of this concept: either "the effect of the availability of leave – for a yet unborn child – on the decision of an individual to have that child," or "the overall effect of leave on the fertility of women across the population." These two conceptualisations clearly correspond to the future-child effect and the total effect. We therefore judge studies identifying the future-child effect or the total effect to have high construct validity, because these studies validly represent the target construct of "the effect of leave on fertility." By contrast, the current-child effect would correspond to the conceptualisation, "the effect of leave on someone who has just had a child, on their subsequent childbearing." We

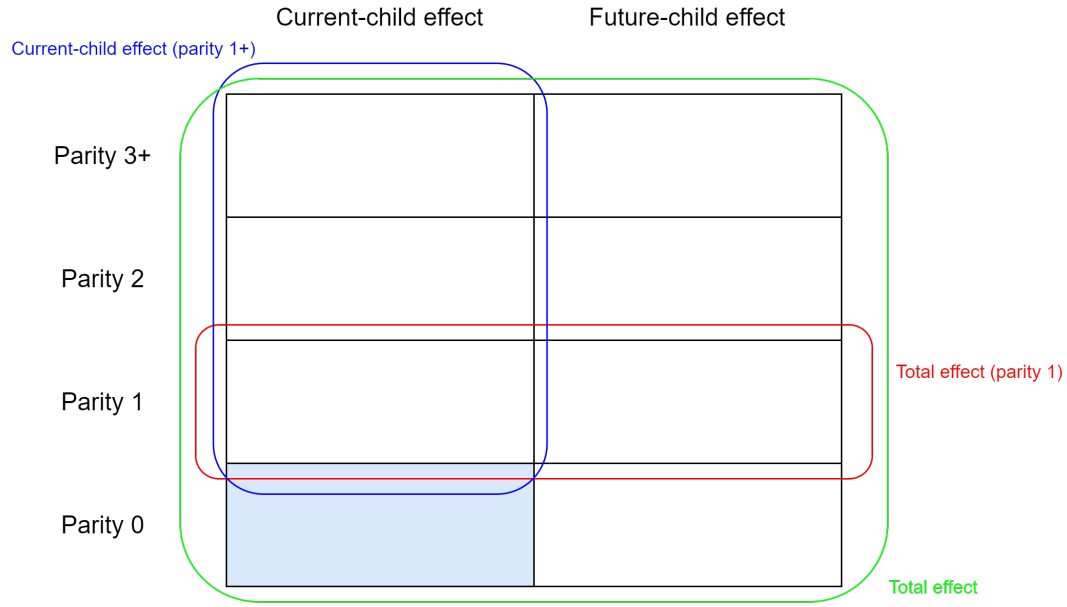


Figure 4.2: Types of effect of a leave policy on fertility, by parity.

Notes: The two columns correspond to the current-child effect and the future-child effect. The rows correspond to the parity of the parent. The blue, red, and green oblongs represent the current-child effect (parity 1+), the total effect (parity 1), and the total effect respectively.

contend that this conceptualisation is not what pro-natal policy-makers mean by "the effect of leave on fertility," and so we judge the current-child effect to have low construct validity. For the remainder of this review, we will divide studies into two categories: current-child effect, and future-child and total effect. Since our objective is to inform pro-natal policy-making, we will give greater weight to future-child and total effect studies.

4.4.2.2 Study designs and identification strategies

The current-child effect can be identified by comparing those giving birth in the weeks or months before the reform, with those giving birth in the weeks or months after the reform. Since these two groups differ in terms of their leave entitlements for the child they just had –and will have the same entitlements for any future child –the current-child effect can be identified by comparing their subsequent fertility over the next several years. Such a study design attempts to approximate a randomised study by arguing that women giving birth shortly before and after the reform are likely to be otherwise similar, meaning that other variables are controlled for. Such a study design is a type of regression discontinuity design (RDD). We define this study design as the "short before-after" design, and illustrate how it works in Figure 4.3(a).

The short before-after design typically requires large administrative datasets, since survey data typically will only have very few women giving birth in the

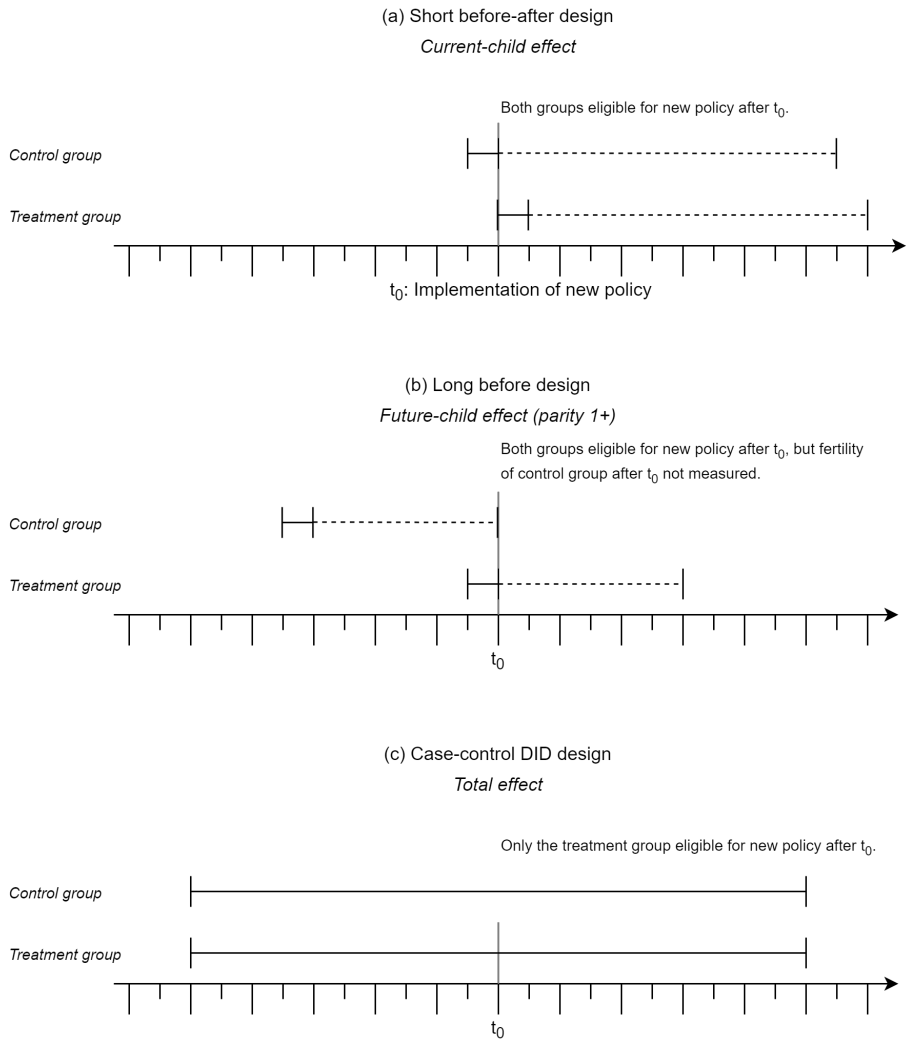


Figure 4.3: Study designs for evaluating the effects of leave on fertility.

Notes: Long check marks on the horizontal axis indicate years, and short check marks indicate half-years. Solid horizontal lines indicate data on births being collected at that time, and dashed horizontal lines indicate that data on births occurring at that time will be collected later.

short periods before and after the reform. Since both groups have access to the new policy after the reform, the short before-after design cannot identify the future-child effect (parity 1+). Furthermore, the short before-after design cannot identify the future-child effect (parity 0), since the analytical sample is restricted to women who have had at least one child.

To identify the future-child effect (for any parity), an alternative study design is required. Lalive and Zweimüller aim to identify the future-child effect (parity 1+) by comparing women who gave birth a month before the reform, with women who gave birth in the same month 3 years before. Both cohorts are therefore entitled to the old policy for the child they have just had, but will have different entitlements if they have another child within the next 3 years. The future-child effect (parity 1+) can then be identified by comparing the fertility of the two

cohorts at the end of their respective three year periods. We define this design as the "long before" study design, and illustrate it in Figure 4.3(b). A major problem with the long before identification strategy is that it assumes there are no systematic differences (that are important for fertility) between the two cohorts. This is a strong assumption: in order for it to be valid, there would have to be no effect of long-term trends in childbearing behaviour. Consequently, estimates for the future-child effect (parity 1+) under this strategy are at risk of bias.

The long before design cannot identify the future-child effect (parity 0) for the same reason that it cannot be identified in the short before-after design: in order to be included in the analytical sample, individuals must have had at least one child. Consequently, neither design can estimate the total effect. In order to estimate the total effect, a case-control study design is required. This means that the new leave policy only becomes available to some women (the case group, who receive the "treatment" of the policy) and not others (the control group, who do not receive treatment). The total effect can then be estimated by comparing the post-reform fertility of the case group and the control group. In a randomised study, individuals would be assigned to treatment randomly; however, this is rarely the case with policy changes, and not the case for any of the studies in this review. In the three case-control studies included in this review, individuals are allocated to treatment by either region, employment status, or income. These studies are problematic in the sense that there are usually pre-existing systematic differences between the case group and the control group, differences which are correlated with either region, employment status, or income. However, a difference-in-differences (DID) approach enables one to approximately control for these differences. We therefore use the term "case-control DID" to refer to this type of study design. A possible implementation of the case-control DID design is illustrated in Figure 4.3(c), using panel data.

While it is possible to identify the total effect using a case-control DID design, the precise effect identified in any given model depend on the sample restrictions imposed by the analyst. For example, in one of the case-control DID models used by Cannonier (2014), the analytical sample is restricted to women who had never given birth at time t_0–4. The outcome variable is a binary indicator, indicating whether the individual had at least one birth between t_0–4 and t_0+17. This method therefore identifies the future-child effect (parity 0). By contrast, in Cannonier' s second model, the analytical sample is restricted to women who had had exactly one birth at time t_0–4. The outcome variable is a binary indicator, indicating whether the individual had at least one more birth between t_0–4 and $t_0 + 17$, and so this model identifies the total effect (parity 1).

4.4.2.3 Categorisation of the studies in the review

Table 4.2 categorises the 23 studies in terms of their study design and the policy effects they identify, and also indicates the sign of the relationship between leave and fertility (final column). Table 4.2 shows that all 6 studies identifying either the future-child effect or a total effect report positive results, whereas the 17 studies that identify the current-child effect report a mixture of negative, null, and positive results. Table 4.2 gives a key for each study, to enable quick referencing.

Table 4.2: Categorisation of the studies in this review, in terms of study design and type of effect identified.

Author(s) (year of publication)	**Study key**	**Year of reform**	**Study design(s)**	**Type of effect on fertility**	**Relationship**	**Generosity**
Ang (2015)	Ang 2006 T0+	2006	Case-control DID	Total effect	Positive	Generous
Cannonier (2014)	Cannonier 1993 FC0	1993	Case-control DID	Future-child effect (parity 0)	Positive	Generous
	Cannonier 1993 T1			Total effect (parity 1)	Positive	Generous
Carneiro, Løken, & Salvanes (2015)	Carneiro 1977 CC1+	1977	Short before-after	Current-child effect (parity 1+)	Null	Ungenerous
Cygan-Rehm (2016)	Cygan-Rehm 2007 CC1+	2007	Short before-after	Current-child effect (parity 1+)	Null	Ungenerous
Dahl et al. (2016)	Dahl 1987 CC1+	1987	Short before-after	Current-child effect (parity 1+)	Null	Ungenerous
	Dahl 1988 CC1+	1988		Current-child effect (parity 1+)	Null	Ungenerous
	Dahl 1989 CC1+	1989		Current-child effect (parity 1+)	Null	Ungenerous
	Dahl 1990 CC1+	1990		Current-child effect (parity 1+)	Null	Ungenerous
	Dahl 1991 CC1+	1991		Current-child effect (parity 1+)	Null	Ungenerous
	Dahl 1992 CC1+	1992		Current-child effect (parity 1+)	Positive	Ungenerous

Table 4.2 cont.

Author(s) (year of publication)	**Study key**	**Year of reform**	**Study design(s)**	**Type of effect on fertility**	**Relationship**	**Generosity**
Duvander, Johansson, & Lappegard (2016)	Duvander 1993 CC1	1993	Short before-after	Current-child effect (parity 1)	Null	Ungenerous
	Duvander 1993 CC2			Current-child effect (parity 2)	Null	Ungenerous
	Duvander 1995 CC1	1995		Current-child effect (parity 1)	Null	Ungenerous
	Duvander 1995 CC2			Current-child effect (parity 2)	Null	Ungenerous
Farré & González (2019)	Farré 2007 CC1+	2007	Short before-after	Current-child effect (parity 1+)	Negative	Ungenerous
Hart, Andersen & Drange (2019)	Hart 2009 CC1+	2009	Short before-after	Current-child effect (parity 1+)	Null	Ungenerous
Lalive & Zweimüller (2009)	Lalive 1990 CC1	1990	Short before-after	Current-child effect (parity 1)	Positive	Generous
	Lalive 1990 FC1		Long before	Future-child effect (parity 1)	Positive	Generous
	Lalive 1996 CC1	1996	Short before-after	Current-child effect (parity 1)	Null	Ungenerous
	Lalive 1996 FC1		Long before	Future-child effect (parity 1)	Positive	Ungenerous
Raute (2019)	Raute 2007 T0+	2007	Case-control DID	Total effect	Positive	Generous

4.4.3 Study quality

We applied the ROBINS-I questionnaire to assess risk of bias (RoB) for all 23 studies. We present our results separately for the current-child effect studies, and for the total effect and future-child effect studies. Before we present our results in Sections 4.4.3.2 and 4.4.3.3, Section 4.4.3.1 explains ROBINS-I in detail, and discusses three considerations that were important for how we adapted ROBINS-I.

4.4.3.1 Application of ROBINS-I

ROBINS-I is a tool developed by medical researchers, intended to be used to evaluate RoB in non-randomised studies of medical interventions (Sterne et al. 2016). This review evaluates the effects of a public policy rather than the effects of a medical intervention, and so special considerations needed to be addressed in order to ensure that ROBINS-I can be applied meaningfully. Public policies differ from medical interventions in terms of their coverage (groups rather than individuals), their causal channels (social rather than biological), and their delivery (administered by state administrators rather than medical professionals). We adapted ROBINS-I by accounting for three special considerations specific to studies of the effect of leave on fertility. We show that ROBINS-I cannot be meaningfully applied without taking these considerations into account, and describe how we account for them. We account for these considerations by altering the meaning and interpretation of the existing bias domains in ROBINS-I.

The first consideration concerns detection bias, and applies to all 23 studies in this review. Detection bias concerns whether a patient in a medical trial is aware that they are receiving the treatment of interest (as opposed to a placebo). Detection bias cannot be controlled for in the studies of the effect of leave on fertility, since everyone receiving parental leave is fully aware that they are receiving parental leave (there is no way of having a placebo for parental leave). We contend that it is inappropriate to compare a non-randomised study of the effect of leave on fertility to a hypothetical, well-conducted RCT in which participants are unaware of their treatment status (as instructed by ROBINS-I). This comparison would be inappropriate since detection bias in studies of leave on fertility cannot realistically be controlled for. Instead, we compare the studies in this review to a hypothetical RCT in which participants are aware of their treatment status, and classify all studies as being at low risk of detection bias.

The second consideration concerns the classification of studies as either "assignment to treatment" studies or "assignment and adherence to treatment" studies, which is done before evaluating the seven domains of bias. This consideration

applies to the four case-control DID studies in this review. As indicated in Section 4.3.1 our review focusses on the effect of leave availability on fertility, rather than the effect of leave use on fertility, and classifies studies of leave availability as "intention to treat" (ITT) (or "assignment to treatment") studies. The problem is that ROBINS-I does not ask questions about time-varying confounding for ITT studies. Time-varying confounding occurs when an intervention can change after the beginning of the study, which risks biasing effects since participants may transition between case and control groups. ROBINS-I does not ask questions about time-varying confounding because individuals in medical ITT studies cannot change their assignment status: assignment is determined by whether or not the doctor prescribes the treatment, and cannot change after the experiment has begun. In the case of parental leave policies however, individuals in the control group can change their assignment status by moving into locations or occupations with more generous entitlements under a new policy (Hong and Sullivan 2016; Nakajima and Tanaka 2014). Such self-selection would risk biasing effect sizes upward, because individuals who might have had children anyway would be sorting themselves into the treatment group. To address this problem, we classified all studies as ITT, but including the potential for time-varying confounding in the list of baseline confounding factors (Domain 1).[4]

The third consideration arises from the distinction between policy "availability" and policy "eligibility". We define an individual as "eligible" for a leave policy if they meet all the criteria necessary to claim entitlements under that policy. We define an individual as having a policy "available" to them if they are either eligible, or if they could become eligible by moving to a different region or entering the workforce. We contend that the effect of policy eligibility is of more interest to policy-makers than the effect of policy availability, since those who are eligible for the policy are the most likely to be affected. In our review, 2 studies evaluated policy availability and 21 studies evaluated policy eligibility. For the application of ROBINS-I, studies of policy availability were assumed to be seeking to estimate the effect of policy eligibility on fertility and therefore to have an additional source of RoB. This additional RoB arises from including individuals who are not eligible for the policy in the "treated" group. Since individuals who are not eligible for the policy are less likely to be affected by the policy than individuals who are eligible, we expect this additional source of RoB to risk biasing study findings towards null. We thus expect studies of leave availability to be

[4]Short before-after studies and long before studies are not at risk of time-varying confounding. In the short before-after design, participants cannot change the fact that they had a birth shortly before or after the reform, so they cannot transition between treatment groups. In the long before design, the case and control groups do not overlap in time, so participants cannot transition between treatment groups.

biased towards not finding any effect, compared with studies of leave eligibility. We accounted for this RoB in Domain 5 (missing data).

4.4.3.2 Current-child effect studies

We identified four major potential sources of RoB in the current-child effect studies. Three are specific to the short before-after study design, and are accounted for by ROBINS-I in Domain 1 (confounding). The fourth is common to all three study designs, and is accounted for in Domain 4 (selection bias).

The first confounding factor is "strategic conception," arising from the fact that if couples know the date of the policy implementation, they can potentially time their conception (and their birth) so that they are eligible for a more generous leave policy. The second confounding factor is "strategic delivery," arising from couples altering the specific day of birth of their child by either bringing forward delivery (e.g. through a caesarean section) or delaying delivery (e.g. through postponing a caesarean section). Strategic conceptions and strategic deliveries may confound study findings because those who are aware of the additional benefits may differ in systematic ways from those who are not aware (e.g. being more educated or having higher incomes). The third confounding factor arises from potential systematic differences between the treatment and control group due to the distribution of births across the seasons. For example, higher-earning couples could time their births around the start of the school year in September, since that would mean their child would have had an extra year of cognitive development compared with their year-group peers that were born in August. We term this confounding factor "birth seasonality."

The fourth potential source of RoB, labelled "other policies," is found in all three study designs. It arises from the simultaneous implementation of additional family policies influencing fertility, benefitting only the treated group. That is, if those giving birth after the reform are eligible for new leave benefits and also for benefits from another new policy (e.g. subsidised childcare), then identification of the effect of the leave policy on fertility will be threatened. This is because there will be two ways in which the case and control groups differ. We account for this RoB in Domain 4 (selection bias).

Figure 4.4 reports the results of running ROBINS-I on the 17 current-child effect studies. All studies were at a low risk of bias in Domains 2-4, 6 and 7, but had higher risks of bias in Domain 1 (confounds) and Domain 5 (missing data). Domain 1 accounts for strategic conceptions, strategic deliveries and birth seasonality. Studies with lower RoB were able to demonstrate that each confound either was not a problem, or was controlled for. Two studies (Farre 2007 CC1+ and Lalive 1996 CC1) were evaluated at critical RoB in Domain 1, because the

studies did not address the risk of strategic conceptions, birth seasonality, or covariate imbalance. All studies were at low RoB in Domain 2 (selection bias), Domain 3 (misclassification bias), and Domain 4 (selection bias), because all studies demonstrated: that their analytical samples were representative of the population exposed to the policy; that study participants were correctly classified; and, that there was no simultaneous introduction of additional family policies. For Domain 5, 14 studies did not directly observe policy eligibility, and instead imputed eligibility from observed variables such as employment history (e.g. Cools 1993 CC1+). Stronger studies mitigated the fact that eligibility was unobserved by demonstrating that their findings were robust to alternative strategies for imputing eligibility. One study (Farre 2007 CC1+) focussed on policy availability, and so was evaluated as being at serious RoB in Domain 5. All studies were evaluated as being at low RoB in Domain 6 because all individuals eligible for a policy are aware that they are eligible, as discussed in the previous section. Lastly, all studies were at low RoB in Domain 7 (outcome reporting bias), because all studies demonstrated their findings to be robust across a range of different model specifications.

Overall, three studies were at low RoB, one study was at moderate RoB, 11 studies were at serious RoB, and two studies were at critical RoB. As recommended by ROBINS-I, we omit the two studies at critical RoB from our synthesis. We thus include 15 studies in our synthesis, giving greater weight to studies displaying low and moderate RoB studies.

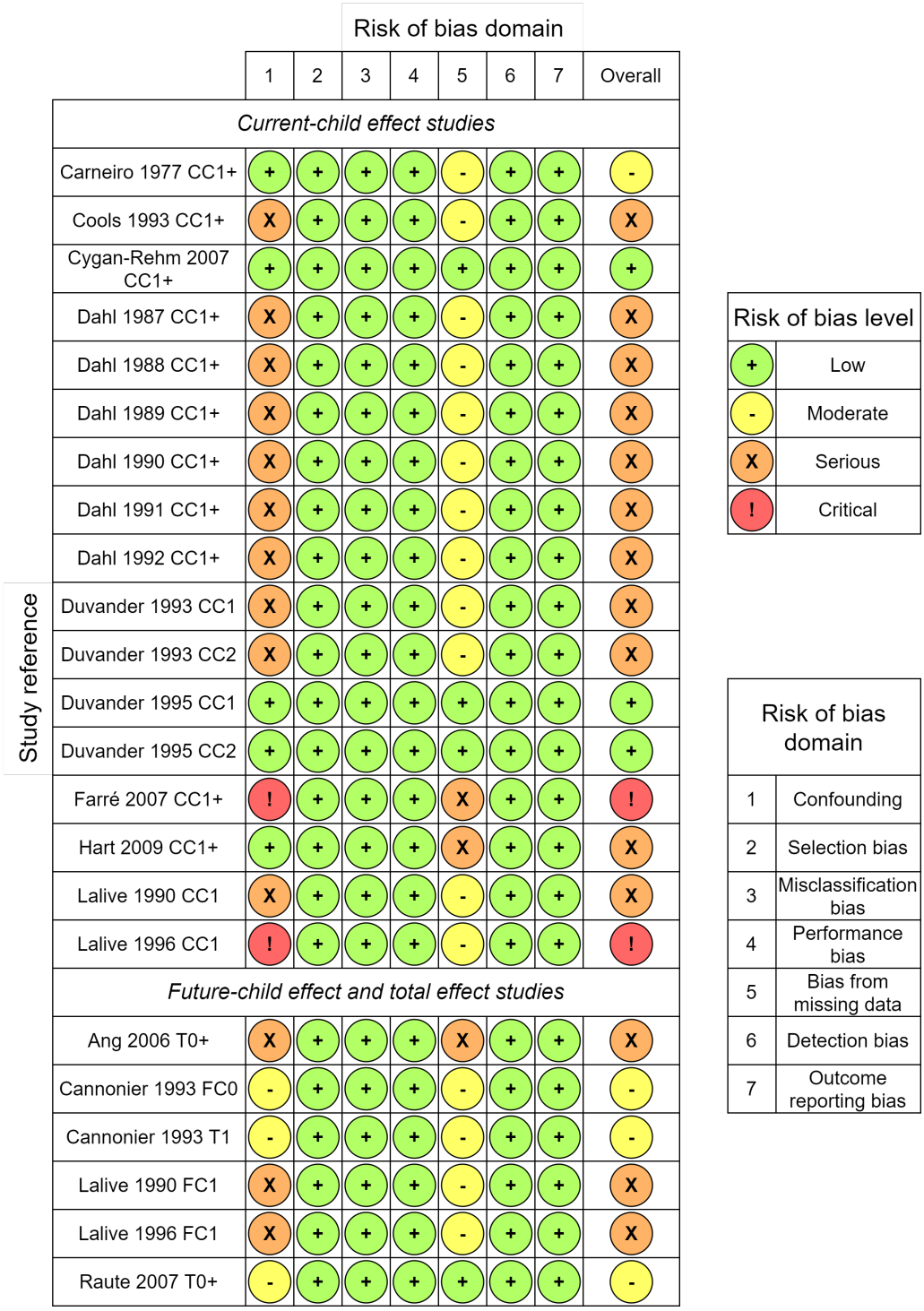


Figure 4.4: ROBINS-I evaluations of included studies.

Notes: Study references are the same as those in the "Study key" column of Table 4.2. Figure adapted from visualisations produced by the R package "robvis" (McGuinness and Higgins 2020).

4.4.3.3 Future-child effect and total effect studies

We identified two potential sources of bias for the future-child effect and total effect studies, one of which applies to the case-control DID study design, and one of which applies to the long before study design. The case-control DID study design is at risk of time-varying confounding, due to the potential for study participants to transition between treatment and control groups after the policy is implemented. We accounted for time-varying confounding in Domain 1 (confounding). The long before study design is also at RoB due to potential long-term trends in fertility and the socioeconomic predictors of fertility. Individuals in the control and treatment groups will have lived through periods of different aggregate fertility, macroeconomic circumstances, and historical events, which may have influenced their fertility choices. This source of RoB is included in Domain 1 (confounding). Both the case-control DID study design and the long before study design are also at RoB due to "other policies."

Of the six studies, four used a case-control DID design and two used a long before design. One of the case-control DID studies (Ang 2006 T0+) evaluated the effect of policy availability, and the five other studies evaluated the effect of policy eligibility. As for the current-child effect studies, Domains 1 and 5 were problematic (see Figure 4.4), with all studies reporting moderate or serious RoB in Domain 1 and four studies reporting moderate RoB in Domain 5.

In Domain 1, the two long before studies (Lalive 1990 FC1 and Lalive 1996 FC1) did not offer evidence that findings were unaffected by long-term trends, and were therefore judged as being at serious risk of bias. The case-control DID study Ang 2006 T0+ was judged to be at serious RoB as it did not control for couples migrating into the region where the new leave policy was introduced. The two case-control DID studies in Cannonier (2014) were evaluated as being at moderate RoB in Domain 1, because while individuals may have self-selected into eligible jobs after the reform, the author provides some evidence to suggest that such self-selection was unlikely to have occurred. In Raute 2007 T0+, the author did not control for self-sorting of individuals who wanted children into the higher-income category. However, since individuals would want to attain higher incomes anyway – and the new policy entitlements may not have provided enough of an extra incentive for individuals to try and increase their income – we judged the RoB in Domain 1 to be moderate, rather than serious. In Domains 2 to 4 all studies were at low RoB, because they demonstrated that their samples were representative, that participants were correctly classified, and that there were no other family policies introduced at the same time.

In Domain 5, the studies in Lalive and Zweimuller (2009) and Cannonier

(2014) were evaluated as being at moderate RoB because eligibility was not directly observed, and was imputed instead. However, study findings were robust to alternative imputation strategies for eligibility, and so these four studies were judged to be at moderate RoB. Ang 2006 T0+ was judged to be at serious RoB in Domain 5 because the study evaluated leave availability rather than eligibility. Raute 2007 T0+ was judged to be at low RoB in Domain 5 because eligibility was directly observed. All studies were classified as being at low RoB in Domain 6 because all participants were aware of whether they were or were not eligible for the policy. All studies were classified as being at low RoB in Domain 7, since a range of alternative models with comparable results were reported. Overall, three studies were at serious RoB, and three studies were at moderate RoB. We therefore include all six studies in our synthesis, giving greater weight to the studies at moderate RoB.

4.4.4 Synthesis of results

4.4.4.1 Current-child effect studies

Almost none of the current-child effect studies found a significant impact of leave on fertility. Out of the 17 current-child effect studies that were evaluated for RoB, two were judged to be at critical RoB overall and are therefore excluded from the synthesis. Two of the 15 remaining studies reported a positive relationship between leave and fertility, and 13 reported no significant relationship. The two studies that found positive results were Dahl 1992 CC1+ and Lalive 1990 CC1, both of which were judged to be at serious overall RoB. Of the 13 studies reporting no significant relationship, three were at low RoB, one was at moderate RoB, and nine were at serious RoB. Dahl 1992 CC1+ found that post-reform mothers had 0.042 more children after 14 years, a finding that was significant at the 10% level. In contrast, Lalive 1990 CC1 found that post-reform mothers were 3.5 percentage points likelier to have a birth up to ten years after the reform, a finding that was significant at the 1% level. In this way, the finding of Dahl 1992 CC1+ is very small and possibly only significant due to sampling error, whereas the finding of Lalive 1990 CC1 is much larger and more clearly significant. All of the current-child effect studies are from Northern and Western Europe: 11 are from Norway, two are from Sweden, one is from Germany, and one is from Austria. Case studies from Norway might appear to dominate our findings; however, excluding the Norwegian studies does not affect our conclusions, since three of the four non-Norwegian studies report null results.

Increases in entitlements under a new leave policy can be conceptualised in terms of whether the policy provides a lot more money or length of leave (i.e.

absolute generosity), and in terms of whether the entitlement increases are large relative to the pre-reform entitlements (i.e. relative generosity). The reform of Lalive 1990 CC1 was both absolutely and relatively generous, doubling the length of leave from 12 months to 24 months and remaining at a flat rate of 340 Euros a month. Moreover, the reform entitled women to automatically renew their leave period if they had another birth within 27.5 months of the previous birth (rather than 15.5 months for pre-reform mothers). The renewal entitlement created a strong incentive for post-reform mothers to have another birth quickly relative to pre-reform mothers (a "speed premium"), since it is biologically feasible to have a birth within 27.5 months of a previous birth, but not feasible within 15.5 months. In contrast, the reform of Dahl 1992 CC1+ was neither absolutely generous nor relative generous. The duration of leave only increased by 3 weeks, from 32 weeks to 35 weeks, and the weekly remuneration rate remained the same.

For 12 of the studies that reported no significant relationship between leave and fertility, increases in leave tended to be between 2-8 weeks with no increase in the rate of remuneration, so the absolute generosity was low. The relative generosity of the reforms in these studies was also low, since all the pre-reform leave duration was at least 18 weeks. The 13th study with null findings evaluated a moderately generous reform (Carneiro 1977 CC1+), in which maternity leave increased from 12 weeks unpaid to 18 weeks at 100% income replacement, plus 1 year unpaid. None of the null effect studies evaluated reforms that introduced speed premiums as in Lalive 1990 CC1. Overall, it seems that generous new entitlements and speed premiums are preconditions for a leave policy having a current-child effect on fertility. However, there is only one study that meets these preconditions, and so we cannot conclude that generous entitlement increases and speed premiums are sufficient for the current-child effect to operate.

4.4.4.2 Future-child effect and total effect studies

All six studies that evaluate either the future-child effect or the total effect find a positive causal impact of leave on fertility. However, we omit one of these studies from the synthesis – Lalive 1996 FC1 – since the authors do not report numerical results for that study.[5] Of the five remaining studies, Ang 2006 T0+, Cannonier 1993 T1 and Raute 2007 T0+ use a case-control DID design to evaluate a total effect; Cannonier 1993 FC0 uses a case-control design to evaluate a future-child effect; and Lalive 1990 FC1 uses a long before design to evaluate a future-child effect. Ang 2006 T0+ and Lalive 1996 FC1 were judged to be at serious RoB, and the other three studies were judged to be at moderate RoB. In terms of effect

[5]The only mention of the results of Lalive 1996 FC1 in Lalive and Zweimüller (2009) is in footnote 32, page 1399.

sizes, Ang 2006 T0+ finds that women benefitting from the reform were 23.53% likelier to have a birth, a finding which is significant at the 1% level. Similarly, Raute 2007 T0+ finds that women benefitting from the reform are 16% likelier to have a birth, a finding which is also significant at the 1% level. Ang 2006 T0+ and Raute 2007 T0+ are the only studies that evaluate the total effect of a leave policy on women at all parities, and therefore provide the best evidence of the impact of leave policies on the aggregate fertility of beneficiaries. The two studies in Cannonier (2014) find the 1993 FMLA reform in the US to make women 5.19 percentage points likelier to have a first birth, and 2.96 percentage points likelier to have a second birth, which are significant at the 5% and 10% level respectively. Similarly, Lalive 1990 FC1 finds women 6.8 percentage points likelier to have a second birth, which is significant at the 5% level. The effect sizes found in these five studies are all quite large, suggesting that leave policies can potentially have a large impact on increasing fertility. In terms of their geographical coverage, 2 studies are from the US, 1 study is from Canada, 1 from Germany, and 1 from Austria.

All five of the future-child effect or total effect studies evaluate reforms that were either absolutely generous or relatively generous. For both of the two total effect parity 0+ studies, the increase in maximum total benefits over the leave period was roughly $20,000 USD. As discussed in the previous section, the 1990 Austrian parental leave reform in Lalive 1990 FC1 doubled the duration of leave from 12 months to 24 months, and kept the same rate of remuneration. The 1993 FMLA reform evaluated by Cannonier (2014) was not absolutely generous, in that it only granted eligible women 12 weeks of unpaid leave. However, prior to the FMLA women did not have any statutory leave entitlements, and so the FMLA represented a large relative increase. Since all five studies evaluate reforms that provided generous increases in benefits, we cannot establish whether an ungenerous policy would also impact fertility through the future-child or total effect. Generosity may be sufficient but not necessary for a policy to impact fertility; alternatively, generosity may be both sufficient and necessary. However, we can conclude that – at a minimum – generosity is sufficient for a new leave policy to increase fertility.

4.5 Discussion

In this section we discuss the implications of our findings for policy and the academic literature and the contribution of our conceptual, methodological and RoB analyses for future research, before identifying some limitations of our review.

Our findings suggest that leave policies are a viable strategy for governments

seeking to increase fertility in low-fertility settings. Moreover, leave policies can be a cost-effective strategy for increasing fertility, compared with other pro-natal policies. When policy-makers discuss the effect of leave on fertility, generally they are referring to either the future-child effect or the total effect – i.e. the extra incentive given to individuals to choose to have a child. The current-child effect is of marginal interest, and it seems that it is overrepresented in the literature since it may be easier to identify empirically. In terms of the academic literature, our analysis and findings might explain why commentators have heretofore evaluated the evidence as "mixed." When treated as an undifferentiated whole, it does seem that the evidence is indeed mixed. However by filtering out studies of the current-child effect from studies of the future-child effect or total effect – and by arguing that the future-child and total effects are the effects important for policy-makers wanting to increase aggregate fertility – we have demonstrated that the evidence for the effect of leave on fertility is entirely supportive. However, a central limitation of this core empirical finding is that our results cannot be generalised to East Asian settings. Since our study only included empirical studies from countries in Europe and North America, it would be problematic to try and transport those findings to settings with very different welfare systems (Esping-Andersen and Billari 2015).

The conceptual, methodological, and RoB sections of this review provide a clear and logical framework for conducting, classifying, and evaluating studies of the effect of leave on fertility. The bifurcation of studies as identifying either the current-child or future-child effect separates studies into two fundamentally different types, meaning that studies evaluating different types of effects cannot be directly compared. Furthermore, our framework classifies studies in terms of the parity of the individuals they analyse. For future empirical research, this framework will enable researchers to better conceptualise and understand the precise effects identified by their studies. This conceptual framework has been developed in order to categorise studies of the effect of leave on fertility, but it could be used to analyse the effects of other pro-natal policies on fertility. Moreover, the study designs associated with evaluating these effects – the short before-after design, the long before design, and the case-control DID design – could also be used for to evaluate the effects of other pro-natal policies.

In terms of RoB, we hope that our analysis in Section 4.4.3 could provide the foundations for a custom-built tool for assessing RoB in studies of public policy evaluations. Historically, systematic review methods in social science have been adapted from pre-existing methods in medical research, and we hope that our analysis will be used in this tradition. Our application of ROBINS-I also shows that case-control DID studies of eligibility are generally at lower RoB than long

before studies or studies of leave availability, suggesting that researchers should choose to use case-control DID evaluations of eligibility where possible.

There are three considerations that may limit the applicability of our findings to other settings. Firstly, the number of studies of the future-child effect or the total effect are quite small, meaning that it is difficult to establish the relative importance of leave generosity. Specifically, all five studies evaluate generous reforms, and so we cannot know whether ungenerous entitlement increases would affect fertility. Secondly, two of the future-child and total effect studies evaluate reforms that were implemented in the recent past (2006 and 2007), and so whether these reforms will have an impact on completed fertility remains to be seen. Lastly, the geographical coverage of these studies is limited to Northern and Western Europe, and North America. This means that generalising these findings to other low-fertility settings (such as Southern Europe, Eastern Europe, and East Asia) may not be appropriate.

4.6 Conclusion

In this review, we sought to examine what the best available evidence showed about the effect of leave on fertility. Our motivation was that fertility is very low in many countries and declining in many more, that national governments devote large resources to increase fertility through family policies, and that most academic commentators argue that the evidence for the effect of leave on fertility is mixed. In conducting the review, we followed a review protocol written prior to searching the literature, focussed only on primary empirical studies with experimental and quasi-experimental designs, and evaluated the quality of studies using the ROBINS-I framework.

We identified 23 studies which examined the impact of leave policies on fertility. In order to understand the seemingly contradictory findings of these studies, we developed a new conceptual framework that enabled us to categorise the studies based on the effects identified, the parity of study participants, and the study designs used to identify effects. This categorisation demonstrated that all of the studies with null or negative findings were only identifying a narrow type of effect of leave on fertility (the current-child effect), an effect which accounts for only a small part of the total effect and is not of interest to policy-makers. Our categorisation also demonstrated that studies identifying a more complete effect (either the future-child effect or the total effect) all had positive and significant findings. Moreover, the effect sizes found in these studies were large, with the probability of a next birth increased by as much as 24%. We therefore reject the contention that the evidence for the effect of leave on fertility is mixed. Rather,

we find that the apparently mixed evidence is simply an artefact of sub-optimal study design. Our results contribute to our understanding of the effect of leave on fertility by showing that different study designs can only identify certain types of effects, and by showing leave can significantly increase fertility when increases in benefits are generous. However, these conclusions must only be limited to European and North American settings, since there were no study results from East Asian countries.

5

Paper 3 – Limited Impact of the Parental Leave Allowance on Fertility in Taiwan

Taiwan has experienced lowest-low fertility since 2003. The 2009 Parental Leave Allowance (PLA) has been the most generous pro-natal policy, dispersing $2.5 billion USD since its inception to women working in the public and private sectors. Whether the PLA affected childbearing remains an open question. Our objective is to empirically assess the causal impact of the PLA on fertility. We use difference-in-differences (DID) multilevel logistic regression models on repeated cross-sectional data from 2003 to 2016. We identify the likelihood of our treatment group of PLA-eligible women before and after 2009 transitioning to a next birth, compared with a control group of PLA-ineligible women before and after 2009. The PLA had no effect on the fertility of women that were always working, but may have increased transitions from first to second births for women who had been in and out of work since marriage. However, this latter result is not robust to alternative specifications. This evidence suggests that parental leave may not be a cost-effective means for increasing fertility in Taiwan or other low fertility countries in East Asia. This study is the first to evaluate the causal effects of pro-natal policies on fertility in East Asian countries.

5.1 Introduction

Low fertility is a major problem in many contemporary developed societies. At the level of the society, long-term low fertility leads to population ageing, a smaller workforce, negative population momentum, and ultimately population decline (Bloom et al. 2010; Caldwell et al. 2002; McDonald 2008). At the individual level low fertility often indicates individuals having fewer children than they desire, and so presents a problem for life satisfaction and emotional well-being (Beaujouan and Berghammer 2019; Casterline and Han 2017; Chen and Yip 2017). Low fertility is particularly acute in Taiwan, which has experienced a Total Fertility Rate (TFR) less than the "lowest-low" level of 1.3 births per woman since 2003 (HFD 2020). In 2010 Taiwan reported a record low TFR of 0.895, the lowest of any major country in history.

Since the mid-2000s, the Taiwanese government has attempted to increase fertility through pro-natal policies, both at the national and local levels (Chen 2012; Hsueh 2018). These policies have included parental leave, childcare subsidies, and tax breaks for parents (ibid.). Coincidentally the Taiwanese TFR increased from 0.895 in 2010 to 1.159 by 2014. Professor Hseuh Cherng-Tay[1] (薛承泰) attributed this increase to the pro-natal policies of the administration in place at that time (HFD 2020; Hsueh 2018). However, to our knowledge there has been no peer-reviewed empirical evidence assessing the effect of these policies on fertility.[2] In their investigation of the impact of housework division on fertility, Cheng and Hsu 2020 argue that pro-natal policies are likely to be most effective when combined with financial aid packages, and when varied by the education of recipients. This is because childbirth intentions were most sensitive to childcare sharing for highly educated women, suggesting that the fertility of these women would be more responsive to fathers taking more responsibility for childcare – something which could be promoted by fathers taking leave. Consequently it is difficult to establish what further pro-natal measures might help increase Taiwanese fertility to sustainable levels.

The Parental Leave Allowance (PLA) – introduced in 2009 – has been the most significant pro-natal policy, both in terms of cost and coverage (Yang 2019). The PLA entitles workers to 6 months of parental leave at 60% of their pay, and roughly $2.5 billion USD has been dispersed through the PLA, to 624k recipients.[3]

[1] Minister for population policy under the KMT administration of 2008-2012.

[2] There exists a blog post discussing the effect of pro-natal policies on fertility in Taiwan, by Professor Wen-Shan Yang (楊文山) (Professor of Sociology, National Taipei University), but there are no formal scientific studies (Yang 2019).

[3] Statistics on uptake of the PLA can be found on the Bureau of Labor Insurance website, https://www.bli.gov.tw/.

A recent systematic review by the present authors found that parental leave has a consistently positive impact on fertility in Western European and North American countries; however, there currently exists no studies of the causal impact of national parental leave policies on fertility in East Asia (Thomas et al. 2019). It is therefore not known whether leave affects fertility in the same way in East Asia, as in western countries.

In this study we aim to evaluate whether the PLA increased fertility in Taiwan, and explore any differences in effects on women at different parities. Our systematic review found that a difference-in-differences (DID) study design is best able to identify causal effects of a leave policy on fertility, among all quasi-experimental study designs (Shadish et al. 2002; Thomas et al. 2019). Using detailed survey data on Taiwanese women's employment and fertility histories from the Women's Marriage, Fertility and Employment Survey (WMFES), we apply a DID study design and identify the causal effect of the PLA on the fertility of working wives. We compare fertility outcomes of working women (who were eligible for the policy) with non-working women (who were not eligible), before and after 2009.

This paper is structured as follows. In the next section, we discuss details of the PLA programme before reviewing relevant theories and empirical evidence of the effect of parental leave on fertility. Section 5.3 discusses the data source used in our analysis before Section 5.4 describes the methods used to evaluate the effect of the PLA. Section 5.5 presents the results. Section 5.6 discusses the implications of our results in terms of policy and research, and a final section concludes.

5.2 Background

5.2.1 The PLA

The PLA was created through an amendment to the Employment Insurance Act (就業保險法) on the 31st of March 2009, and came into effect on the 1st of May 2009.[4] The PLA entitles all workers participating in Employment Insurance (i.e. private-sector workers) to receive 6 months of pay, at 60% of their average monthly insurance salary for 6 months of leave. While Employment Insurance only applies to private-sector workers, the PLA was also introduced for government workers at the same time. Prior to 2009, parents had been entitled to 2 years leave without

[4]The legislative history of the Employment Insurance Act, the full text of the Act, and associated legislation, can be found online on the "Legislative Yuan legal system" website (https://lis.ly.gov.tw/lglawc/lglawkm).

pay. The PLA entitled workers to be paid for 6 months of that 2-year period. The PLA can be taken by both parents at any time while a child is below the age of 3. However, parents cannot take the PLA simultaneously. From 2009-2019, $68.5 billion NTD has been disbursed through the PLA, about $2.5 billion USD. There have been roughly 624k recipients, receiving an average of $4,000 USD over a 6 month period.

5.2.2 Leave use versus leave eligibility

A key distinction is the use of parental leave versus being eligible for parental leave. Here, we are interested in the effect of leave eligibility on fertility, rather than the effect of leave use on fertility. Conceptually, this is because evaluating the effect of leave use is problematic. Leave is only used after a women has a child. Leave is used after giving birth, and leave use does not lead to a birth. It is knowing that leave will be available after giving birth which may result in an individual or family decision to give birth.

In experimental design terminology, we are evaluating an "intention to treat" (ITT) effect. That means that we are interested in the effect of a policy being available to an individual, regardless of whether they actually decide to take leave and receive payment if they have a birth. We choose to evaluate the ITT effect rather than the "per protocol" effect (i.e. the effect of using leave on fertility) because there is less risk of bias in evaluating ITT effects (Higgins and Green 2008). Individuals may choose to have a child because they know the PLA is available if they need it, but then not actually use it (and instead return to work) after giving birth. This effect would be captured by an ITT study design, but not by a per-protocol design.

5.2.3 The effect of leave on fertility

Parental leave can increase fertility by altering the economics of childbearing and childrearing, by enabling mothers to remain in employment, and by altering the gendered distribution of domestic labour (Becker 1973; Ermisch 2003; McDonald 2008). In economic terms, parental leave pays parents to care for their children, labour which would have been unpaid of the absence of leave. For women who want to remain in work, leave enables them to take time off work without losing their job. Parental leave can also make childbearing more attractive for women if their male partners will also take leave, and increase their participation in domestic labour (Baum and Ruhm 2016; Pronzato 2009; Tamm 2019). The gendered distribution of domestic labour has become a particularly salient contextual feature of childbearing in Taiwan (Raymo et al. 2015). Since 2016, the WMFES has

started asking husbands (as well as wives) about their participation in domestic labour.

In order to statistically identify effects of parental leave policies in a way that is useful for policy-makers, it is important to adopt an appropriate study design. The authors of this study conducted a systematic review into the effect of parental leave on fertility, focusing only on quasi-experimental studies that could reliably identify causal effects. We constructed a novel conceptual framework, which classified studies in terms of whether they identified a "current-child effect," a "future-child effect," or a total effect of a leave policy on fertility. The future-child effect refers to the effect of a leave policy on couples' decisions to have a child, whereas the current-child effect refers to the effect of taking leave for a child just born, on couples' decisions to have another child. The total effect is the sum of these two effects. The current-child effect is however of marginal policy interest, but often the focus of empirical studies as it is easier to design experiment for its statistical identification. By contrast, future-child effect studies have a greater potential to inform pro-natal policy development, since they can more accurately capture "the actual effect of leave on fertility." When all the findings from quasi-experimental studies based future-child effect and total effect studies, the existing evidence consistently showed that leave increases fertility. Yet, the number of final studies is small and focused on leave policies in Western Europe and North America. Whether or not leave policies increase fertility in other low-fertility regions – including East Asia – remains an open question.

5.3 Data

Next we first introduce our data source, and describe the employment information used to identify the group of PLA-eligible women.

5.3.1 The WMFES

Our data source is the Women's Marriage, Fertility and Employment Survey (WMFES). The WMFES is a repeated cross-sectional survey, with each survey comprising of a completely new set of individuals (i.e. the same individuals are not followed over time). The WMFES is conducted by the Taiwanese government, and about 20,000 women aged 15 and over are interviewed in each survey. In this study we use the last 5 surveys of the WMFES, conducted in 2003, 2006, 2010, 2013, and 2016. Key information for our purposes includes birth month and birth year for up to 4 births (the first three births and the most recent birth), wife's age at first marriage, and details of employment before and after marriage and childbearing.

The WMFES enables us to re-construct the respondent's marriage, childbearing, and employment histories. The WMFES does not include information on whether the respondent has even taken parental leave. The use of the PLA is not observed. We restrict our analytical sample to married women, because over 95% births in Taiwan occur within marriages: only 3-4% of births are had by unmarried women (Hsueh 2018). More specifically, we restricted our analysis to births had by women who were married at the time of birth, i.e. ignoring pre-marital births had by women who were married by the time of the survey. We further restrict our analytical sample to women in the childbearing ages of 15-49, who were married at least 4 years before the survey year and were still married at the time of the survey. We also restricted the sample to women who had had no more than 3 births, 4 years before the survey year. We did this because the WMFES only records birth data for 4 births, and because our outcome of interest is whether the respondent had a child in the 3 years preceding the survey year (we explain the rationale of this decision in Section 5.4.4). Our final sample included 34,551 women from an initial sample of 143,377.

5.3.2 Employment histories and leave eligibility

Individuals are only eligible for the PLA if they are in work, and so it is important to establish the employment histories of women in the WMFES. The WMFES asks married women about their employment history up to the time of the survey. The ten possible responses are listed in Table 5.1, along with the proportion of married women giving each response. This information is useful for our purposes, because we are interested in whether eligibility for the PLA caused women to be more likely to have a child in the three years preceding the survey, which we refer to as the "focal period" (our reasons for selecting this outcome variable are explained in Section 5.4.4). An important variable for predicting childbearing in the focal period is the respondent's parity in the year before the focal period – henceforth when we refer to a respondent's "parity," we mean their parity in the year before the focal period, not their parity in the survey year.

The ten employment history categories in Table 5.1 enable us to infer which respondents were or were not working during the focal period. For example, we can infer that Group (1) (wives who have always been working) must have been working, and that Groups (4) and (10) (wives who have not worked at least since marriage) must not have been working. Having determined which respondents were working in the focal period, we can then apply the restriction that those respondents were also employed either in the public sector or private sector (i.e. not self-employed, an employer, or an unpaid family worker), and

thereby determine who was eligible for the PLA in the focal period.

Table 5.1: Employment history categories.

Group	Employment history	Proportion of wives	Employment status in focal period
(1)	Always working	0.340	Working
(2)	Quit work at marriage and have a job now	0.107	Likely working
(3)	Quit work at marriage, had a job after but don't have a job now	0.016	Likely working
(4)	Quit work at marriage, have not had a job after until now	0.140	Not working
(5)	Quit work at birth and have a job now	0.074	Parity 0: working; Parity 1+: likely working
(6)	Quit work at birth	0.087	Parity 0: working; Parity 1+: likely not working
(7)	Quit work for other reasons and have a job now	0.039	Unclear
(8)	Quit work for other reasons	0.049	Unclear
(9)	No job before marriage, but had a job after marriage until now	0.073	Unclear
(10)	Never had a job	0.074	Not working

Notes: The 'Proportion of wives' column gives the proportion of all married women (aged 15-49) in each group across all the surveys. In the "Employment status in focal period" column, women within a given group may have different employment statuses depending on their parity in the year before the focal period. N=34,551.

While we can infer that women in Groups (1), (4) and (10) were working in the focal period, determining the employment status (in the focal period) of women in the other groups is more difficult. For example, Group (2) women quit work at marriage and have a job in the survey year; this information alone is insufficient to determine whether they were working or not in the focal period. However, we can use information from other questions in the survey to estimate the employment status of these other women. The fourth column in Table 5.1 classifies each group as working in the focal period, likely working, not working, likely not working, or whether their work status during the focal period is unclear. Where appropriate, we also specify any within-group differences in focal period employment status due to parity in the year before the focal period. For example, Group (6) women (quit work at birth) who were parity 0 in the year before the focal period had not yet had any children, and so they could not yet have quit

due to a birth. Appendix C.1 discusses how we estimate the employment status in the focal period for each group.

We are interested in the effect of the PLA on women who would have been eligible in the focal period (Group (1) and parity 0 Group (6)), but we are also interested in the effect of the PLA on women who may have been eligible (Groups (2), (3), (5), (7), (8), and (9)). Women who may have been eligible represent a class of women who were neither purely in work or purely out of work since marriage, and so their childbearing may be more responsive to the PLA, which enables them to balance work and childbearing. We therefore combine the ten groups in Table 5.1 into the three groups in Table 5.2, representing women who were always working at least since marriage (Group (A)), women who were occasionally working since marriage (Group (B)), and women who were not working since marriage (Group (C)). We also create these three groups in order to increase statistical power and enable robust estimation of our models.

Table 5.2: Combined employment history groups.

Combined group	Groups	Description	Proportion of wives
(A)	(1), (5) parity 0, (6) parity 0	Always working at least since marriage	0.413
(B)	(2), (3), (5) parity 1+, (7), (8), (9)	Occasionally working since marriage	0.372
(C)	(4), (10)	Out of work at least since marriage	0.214

Notes: For definitions of Groups (1)-(10) see Table 5.1. The 'Proportion of wives' column gives the proportion of all married women (aged 15-49) in each group across all the surveys. We do not include Group (6) parity 1+ women in Group (B) because they were likely not working during the focal period (see Appendix C.1). N=34,551.

To estimate the effect of the PLA, we will compare women in Group (A) with women in Group (C), and we will also compare women in Group (B) with Group (C). Because Group (B) is a highly heterogeneous group, we will also separately estimate models on each of its sub-groups (i.e. (2), (3) etc.), as well as other combinations of Group (B)'s subgroups (e.g. the effect on women in Groups (2), (3), and (5)).

5.4 Method

5.4.1 Differences-in-differences

To evaluate the causal impact of the PLA on fertility, we use a difference-in-differences (DID) estimation strategy and regression models, a standard approach for causal inference in econometrics (Angrist and Pischke 2009). Our systematic review found DID to be an established approach for evaluating the causal impact of leave policies on fertility (e.g. Ang 2015; Cannonier 2014; Raute 2019). Moreover, DID is able to identify both the future-child and total effects, and to identify the effect of a parental leave policy on individuals of all parities.

DID is applied to quasi-experimental studies, in which one group receives "treatment" and is compared to a control group which does not receive treatment. In quasi-experimental studies, we are interested in the effect of a treatment on a particular outcome. If individuals are allocated to the treatment and control groups randomly, then the two groups will be roughly comparable in terms of other predictors of the outcome. With random allocation, the effect of the treatment can be estimated by directly comparing outcomes in the treatment and control groups. However, if individuals are allocated to the treatment and control groups non-randomly, then a direct comparison of outcomes will give a biased estimate of the treatment effect. This is because the two groups may be systematically different in terms of other factors which may also affect the outcome. DID enables us to capture treatment effects in non-randomised studies.

DID aims to identify treatment effects by comparing trends in the outcome for the treatment and control groups over time. In the simplest case, there is one 'before' period t_0 (i.e. before the treatment is applied) and one 'after' period t_1. Based on these categories, four sub-groups can be identified: treated before, treated after, non-treated before, and non-treated after. The 'treated before' sub-group does not receive treatment – the individuals in this group are just people who are similar in some way to those who do receive treatment at time t_1.

In our case, DID involves evaluating whether the difference in fertility between (occasionally) working women and non-working women changed after the PLA was implemented in 2009. Specifically, we want to evaluate the effect of the PLA on wives who have always been working since marriage (Group (A) in Table 5.2), and the effect of the PLA on wives who have been occasionally working since marriage (Group (B)). To estimate the average treatment effect of the PLA, we compare the "treated" Groups (A) and (B), against the "control" Group (C) (i.e. women who have not worked at all since marriage). Separate DID models are estimated comparing: (1) Group (A) with Group (C), and (2) Group (B) and Group (C). Any deviation in average fertility trends of these groups after 2009 can be attributed to a causal impact of the PLA, assuming that differences in

fertility between (occasionally) working women and non-working women would have remained the same without the PLA as discussed below.

5.4.2 DID assumptions

There are two key assumptions required for DID to be valid: the common trends assumption (CT), and the common support assumption (COSU) (Angrist and Pischke 2009; Keng and Sheu 2011; Lechner 2010). CT requires that time trends for the treatment and control groups are parallel prior to the introduction of the policy. In our case, CT states that the difference in fertility between (occasionally) working wives and non-working wives would have remained constant over time if the PLA did not have an impact on fertility. Graphically this means that the fertility time-trends of these two groups should be parallel to each other, prior to 2009. CT cannot be tested directly, since the fertility of (occasionally) working wives in the absence of the PLA after 2009 is not observed. However, CT can be graphically evaluated by looking at whether trends in fertility between the treated and non-treated groups were roughly parallel prior to 2009 (Wing et al. 2018). CT can also be evaluated by estimating two regression modes: a base model and an extended model; that is, the base model including an additive interaction term between group membership (either working or non-working) and a time index (ibid.). If the model results are robust to changing the specification, then CT is valid.

COSU concerns the distributions (the "support") of other predictors (i.e. other than the DID variables) of the outcome variable in the treatment and control groups. COSU states that these distributions within groups must remain stable over time. The distribution of covariates in the treated-before and treated-after subgroups must be similar. Similarly the distribution of covariates in the control before and control after subgroups must be similar (Lechner 2010). In our case, this means that the distributions of variables such as age and education must be similar within the treatment and control groups. For example, if it was the case that the post-2009 working women in our data were on average much older than the pre-2009 working women, then it would not be possible for the DID model to separate the effect of the PLA on fertility from the effect of age on fertility. If COSU is violated, then this violation can be controlled for by including other predictors of fertility in the DID model specification (Stuart et al. 2014). We therefore estimate a DID regressions without covariates and with covariates (Section 5.4.5).

5.4.3 Multilevel modelling

To control for pro-natal policy variation in the 20 local government areas, we use multilevel regression modelling. In multilevel models the parameter(s) of interest are modelled as random variables rather than constants (Gelman and Hill 2007). In our case, we allow the intercept term in the regression model to vary by local government area, meaning that we control for any potential effect of unobserved variation in local government policies. Multilevel models have the advantage that they can estimate models with small numbers of individuals in each group. For example, it may be the case that there are very few observations of working women in the 2013 survey, at parity 0, and who live in Taiwan. In this case, it would not be possible to estimate a classical regression model with dummy variables for county residence.

5.4.4 Variables

Our outcome variable is a binary variable indicating whether the respondent had one or more births in the three full years prior to the survey year. For example, the outcome variable for women from the 2016 survey indicates whether they had one or more births from 2013-2015 inclusive. We refer to this period as the "focal period." We selected this measure of fertility for three reasons. Firstly, it captures fertility behaviour in every year from 2000-2015 inclusive (except 2006). Secondly, it averages out birth fluctuations due to the Chinese calendar (numbers of births drop in the Year of the Tiger and spike in the Year of the Dragon). Thirdly, it omits childbearing information from years in which a given survey was conducted (i.e. any information in the 2016 survey on childbearing in 2016 is omitted) because some respondents may have had births later in the year, after the survey was conducted.

Data for covariates of fertility is taken from four years prior to the survey year, i.e. one year before the 3 year focal period. For example, for women in the 2016 survey, we record covariates in 2012. We do this because in order for there to be a causal impact of a covariate on fertility, that covariate needs to be measured before fertility: the cause has to happen before the effect (Shadish et al. 2002). Our covariates are: wife's parity, wife's age, binary indicators for wife's highest level of education, and binary indicators for husband's highest level of education. We assume that respondents' highest level of education in the survey year is unchanged from four years previously. This is a reasonable assumption as most respondents will have completed their education. Unfortunately, we are unable to register key variables of interest that may be important for prior childbearing. Specifically, we are unable to measure potentially relevant covariates of fertility

such as income and parental co-residence.

Table 5.3: Descriptive statistics.

Variable	Group (A)		Group (B)		Group (C)	
	t_0	t_1	t_0	t_1	t_0	t_1
Birth in focal period	0.144	0.127	0.098	0.127	0.151	0.152
Parity before focal period						
– Parity 0	0.095	0.127	0.045	0.065	0.077	0.106
– Parity 1	0.207	0.244	0.156	0.218	0.177	0.233
– Parity 2	0.450	0.472	0.458	0.494	0.416	0.438
– Parity 3	0.212	0.141	0.285	0.197	0.270	0.193
Mean age of wife	35.1	36.1	36.3	37.1	34.9	36.0
Wife's education						
– Below senior high school	0.255	0.144	0.355	0.185	0.452	0.283
– Senior high school	0.571	0.558	0.574	0.637	0.505	0.601
– Tertiary education	0.174	0.298	0.071	0.178	0.043	0.116
Husband's education						
– Below senior high school	0.310	0.228	0.376	0.251	0.457	0.327
– Senior high school	0.508	0.510	0.518	0.542	0.467	0.533
– Tertiary education	0.182	0.262	0.106	0.207	0.076	0.140
n	5,147	9,759	9,331	4,044	5,433	2,228

Notes: For each of the three groups, observations are split between t_0 (before the PLA) and t_1 (after the PLA). Proportions of respondents at different parities is measured in the year before the 3 year focal period, as is the mean age of the wife.

5.4.5 Model specifications

Our outcome variable y_i is a binary indicator for whether the respondent gave birth to one or more children in the three years prior to the survey, and so we use logistic regression. We run separate sets of regression models for Group (A) (always working women) and Group (B) (occasionally working women). Our simple (i.e. without covariates) multilevel DID models are of the form shown in Equation 5.1.

$$ln\left(\frac{Pr(y_i = 1|\boldsymbol{x}_i)}{Pr(y_i = 0|\boldsymbol{x}_i)}\right) = \alpha_{j,k[i]} + time_i\beta_1 + treat_i\beta_2 + DID_i\beta_3 + \epsilon_i. \quad (5.1)$$

Equation 5.1 is a non-nested, varying-intercept multilevel model, meaning that the intercept $\alpha_{j,k[i]}$ is a random variable, depending on the county j and the

parity k, which is determined by the respondent i (Gelman and Hill 2007). The variables *time* and *treat* are binary indicators, corresponding to whether the 3 year focal period is after 2009 and whether the respondent is in the treatment group. The DID variable is a interaction term between *time* and *treat*, and the value of the parameter β_3 is what we are primarily interested in finding. Similar varying-intercept logistic DID model specifications can be found in Masiano et al. (2019) and Tang et al. (2017).

Assuming that other covariates of fertility are stable over time (i.e. assuming COSU), β_3 will validly capture the effect of the PLA on (occasionally) working women. We test the validity of COSU by estimating the same models with covariates, and seeing whether results change. The DID model with covariates acts as a control for the violation of COSU, enabling unbiased estimation of β_3. The form of models with covariates is specified in Equation 5.2.

$$\begin{aligned} ln\left(\frac{Pr(y_i=1|\boldsymbol{x}_i)}{Pr(y_i=0|\boldsymbol{x}_i)}\right) = \alpha_{k[i]} + time_i\beta_1 + treat_i\beta_2 + DID_i\beta_3 + \\ + age_i\beta_4 + edu\_shs_i\beta_5 + edu\_uni\beta_6 \\ + h\_edu\_shs_i\beta_7 + h\_edu\_uni\beta_8 + \epsilon_i. \end{aligned} \quad (5.2)$$

In Equation 5.2, age_i is the wife's age in the survey year, and the variables containing '*edu*' comprise binary indicators for whether the wife or husband have at most a senior high school or higher education. In Equation 5.2 the parameter $\alpha_{k[i]}$ only varies by parity and not county, because with varying counties there are too many coefficients for the model to be estimated.

Equations 5.1 and 5.2 enable identification of the effect of the PLA on fertility, controlling for differences in intercept due to parity or county. However, we are also interested in whether the PLA impacted women of different parities differently, i.e. whether the β_3 variable has different values for women at different parities. We therefore run separate models corresponding to Equations 5.1 and 5.2 for women at parities of 0, 1, 2 and 3. In the parity-specific version of Equation 5.1, the intercept is $\alpha_{j[i]}$ and only varies by county. In the parity-specific version of Equation 5.2, the intercept is α and does not vary.

5.5 Results

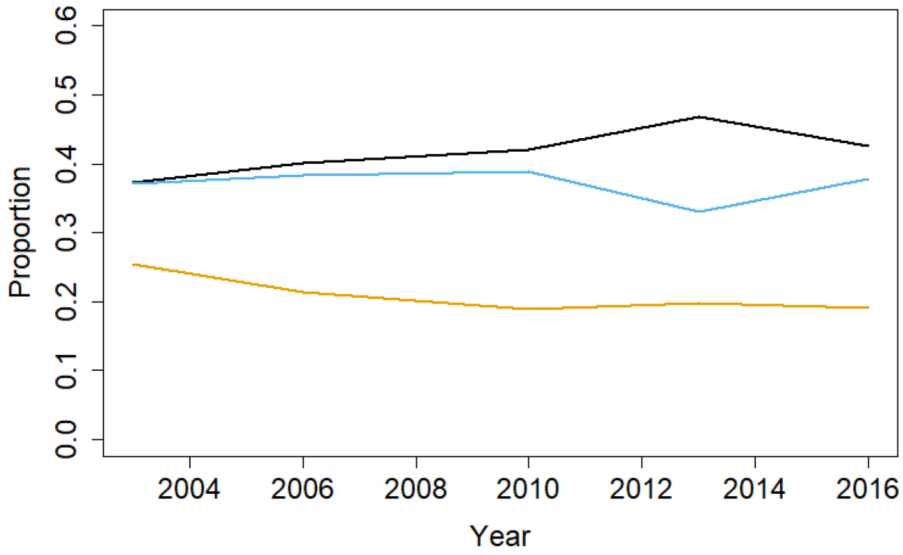
5.5.1 Graphical evidence

Figure 5.1 Panel (a) shows the proportions of women in Groups (A), (B) and (C) over time, and Figure 5.1 Panel (b) shows the proportions of women having a

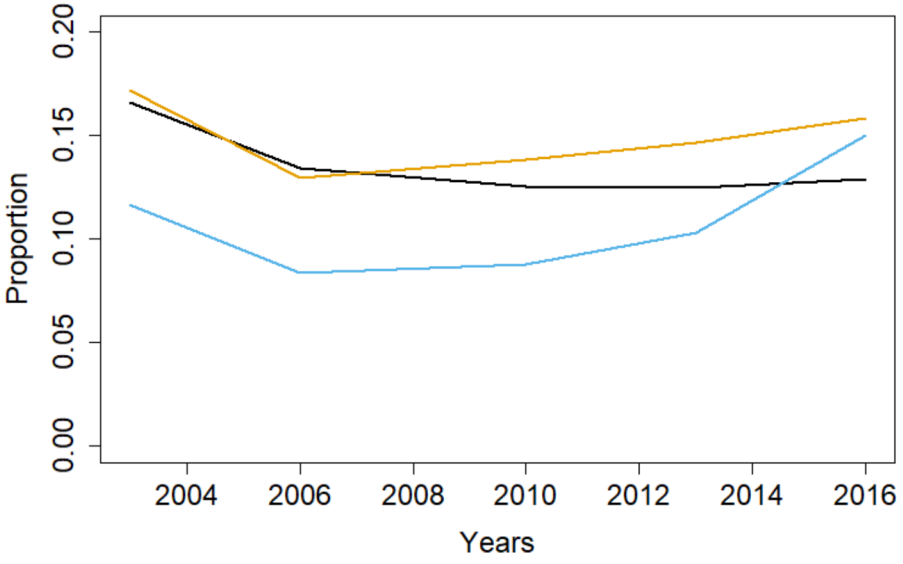
birth in the 3 year focal period over time. Trends for all 10 subgroups are shown in Figure C1 of Appendix C.2. Panel (a) indicates that the proportions in each group remained roughly stable, apart from in the 2013 survey where the proportion in Group (A) increased sharply and the proportion in Group (B) dipped. Panel (b) indicates that fertility trends across groups were roughly parallel from 2003-2010, and especially between Groups (B) and (C). Panel (b) therefore provides support for the validity of the Common Trends assumption. After 2010, the trend in Group (A) does not change relative to the trend in Group (C), suggesting that there was no effect of the PLA on women that were always working. By contrast, the time trend for Group (B) after 2010 curves upward relative to Group (C), suggesting a positive impact of the PLA on the fertility of occasionally working women.

Figure 5.2 breaks down the graphs in Figure 5.1 by parity. Panel (a) indicates that the proportion of women in Group (C) remained relatively stable across all surveys for all parities, and that changes in Group (A) were mirrored by changes in Group (B). Panel (b) suggests that the Common Trends assumption is valid for parities 0, 1 and 2, but not for parity 3. In terms of post-reform trends, Panel (b) indicates that fertility in Group (A) not only did not increase relative to Group (C), but possibly even declined for parities 0 and 2. The fertility trend for Group (B) appears to roughly follow the Group (C) trend for parities 0 and 2 after 2010, suggesting no impact of the PLA on women of these parities. However, the Panel (b) graph for parity 1 women indicates a post-reform fertility increase for Group (B) relative to Group (C), suggesting that the PLA had a positive impact on the second births of occasionally working women.

(a) Proportion of women in each of 3 employment history groups.



(b) Proportion of women having a birth in the 3 year focal period, by 3 employment history groups.



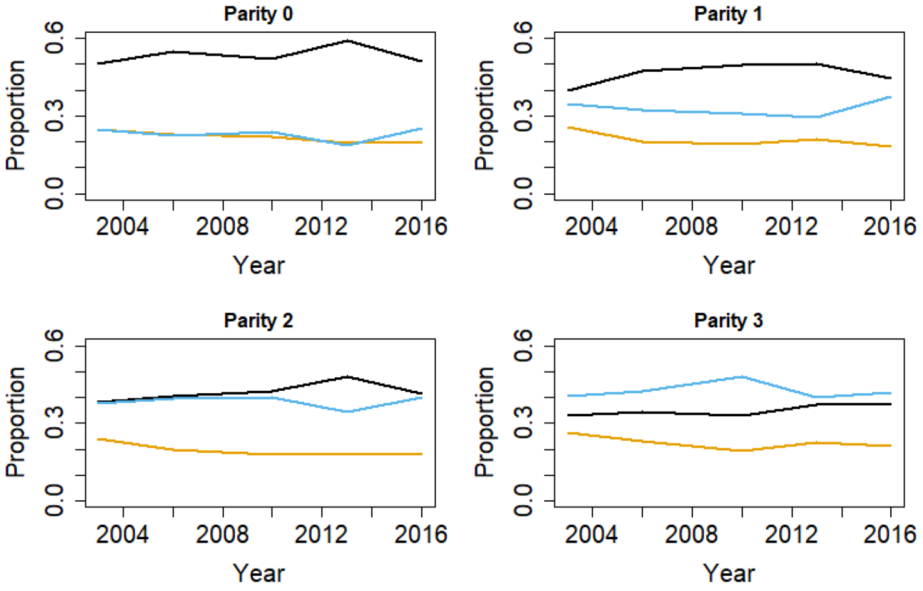
Group (A): Always working at least since marriage

Group (B): Occasionally working since marriage

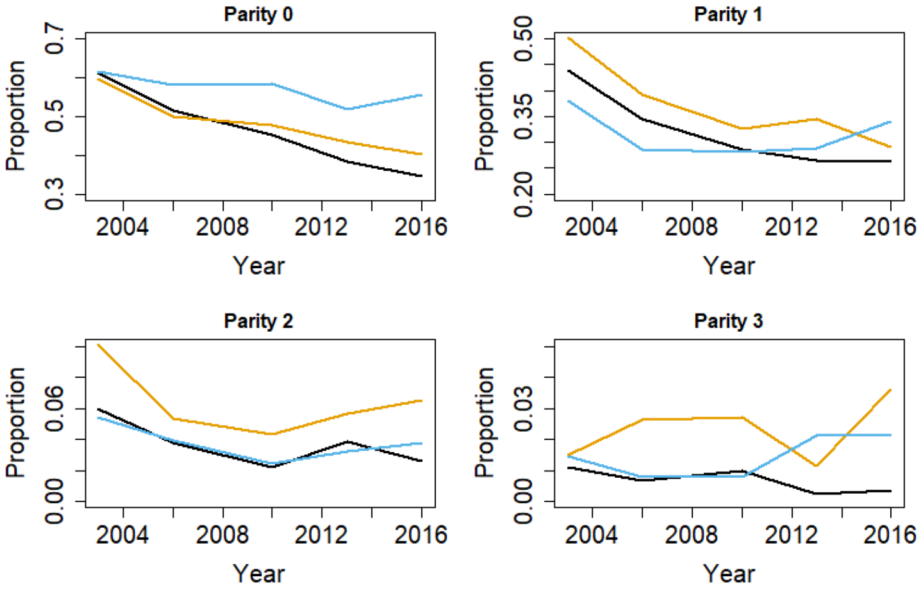
Group (C): Not working at least since marriage

Figure 5.1: Proportions and fertility of wives in three employment history groups.

(a) Proportion of women in each of 3 employment history groups, by parity.



(b) Proportion of women having a birth in the 3 year focal period, by 3 employment history groups, by parity.



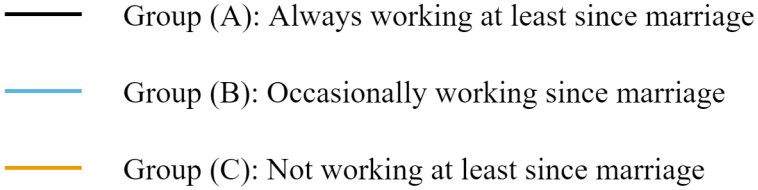


Figure 5.2: Proportions and fertility of wives in three employment history groups, by parity.

Notes: Parity in this case refers to the number of births had by the respondent 4 years before the survey year, i.e. 1 year before the 3 year focal period. Note that the vertical axes in panel (b) all have different ranges.

5.5.2 Always working women

The results for models corresponding to Equations 5.1 and 5.2 are shown in Table 5.4 in Panels (a) and (b), respectively. Specifically, Table 5.4 reports the value of the coefficient of the DID_i variable (β_3) in Equations 5.1 and 5.2, as well as the exponent of β_3. Full model results are reported in Tables C1 and C2 in Appendix C.3. $exp(\beta_3)$ is the odds ratio of β_3, and can be understood as the causal impact of the PLA on the likelihood of having a birth. A value of $exp(\beta_3)$ below 1 indicates a negative impact, a value of 1 indicates no impact, and a value above 1 indicates a positive impact on fertility.

Table 5.4: Always working women (Group (A)) model results.

	All	Parity 0	Parity 1	Parity 2	Parity 3
(a) Without covariates					
β_3	0.04	−0.12	0.11	−0.03	−1.61
	(0.10)	(0.19)	(0.14)	(0.22)	(1.10)
$exp(\beta_3)$	1.05	0.87	1.12	0.97	0.20
(b) With covariates					
β_3	−0.05	0.22	0.07	−0.07	−1.78
	(0.11)	(0.15)	(0.17)	(0.23)	(1.11)
$exp(\beta_3)$	0.95	1.25	1.07	0.93	0.17
n	19,518	2,325	4,243	8,855	4,095

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $+p < 0.1$.

Notes: β_3 is the coefficient of the DID_i variable in Equations 5.1 and 5.2. $exp(\beta_3)$ is the natural exponent of β_3.

None of the values of β_3 in Table 5.4 are statistically significant, even at the 10% level. Moreover, most of the coefficient estimates are smaller than the corresponding standard errors. This indicates that there was no causal impact of the PLA on the fertility of working women. This finding is also true of women at specific parities. Running the same models again, but restricting treated women to those that were working in the public or private sectors, does not substantially change the results.

5.5.3 Occasionally working women

The results for all occasionally working combined (i.e. Group (B)) are shown in Table 5.5, and full model results are shown in Tables C3 and C4 in Appendix C.3. The structure of Table 5.5 is the same as Table 5.4. Similar to Group (A), the results for Group (B) indicate that there was no impact of the PLA on occasionally working women overall. The results are substantially unchanged by

restricting the treatment group to those who were working in the public sector or private sector. To test the Common Trends assumption, we run the same models but include an interaction term between survey year and the treatment group binary indicator. The results are substantially unchanged, indicating that the CT assumption is valid.

Table 5.5: Occasionally working women (Group (B)) model results.

	All	Parity 0	Parity 1	Parity 2	Parity 3
(a) Without covariates					
β_3	0.07	−0.09	0.12	−0.04	0.49
	(0.11)	(0.25)	(0.16)	(0.23)	(0.52)
$exp(\beta_3)$	1.07	0.91	1.13	0.96	1.63
(b) With covariates					
β_3	0.00	−0.08	0.08	−0.12	0.24
	(0.13)	(0.30)	(0.18)	(0.24)	(0.55)
$exp(\beta_3)$	1.00	0.92	1.08	0.89	1.27
n	19,439	1,250	3,676	9,265	5,248

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{+}p < 0.1$.

Notes: β_3 is the coefficient of the DID_i variable in Equations 5.1 and 5.2. $exp(\beta_3)$ is the natural exponent of β_3.

We tested the robustness of the results for Group (B) by running similar regressions on each of the subgroups of Group (B), both individually and in various combinations. None of the subgroups individually reported significant results for any parity-specific models with covariates. Of the combined groups, some combinations gave statistically significant and positive values for β_3, for all women and women at parity 1. Generally, the inclusion of covariates reduced the magnitude and statistical significance of the results for all women and parity 1 women, indicating that the Common Support assumption may not be justified. We therefore prioritise the results from models that include covariates.

Group combination models that reported significant results all excluded Group (9) women (always working since marriage), but included most of the other groups in Group (B). For all women, values of β_3 in the models with covariates ranged between 0.25-0.29, and were significant either at the 10% or 5% level. For parity 1 women, the values ranged between 0.31-0.36, again significant at either the 10% or 5% level. The odds ratios of the mid-points of these ranges (0.27 and 0.34) are 1.31 and 1.40 respectively. To understand the meaning of these numbers, suppose that the pre-reform probability of an occasionally working woman having a birth was 0.10 (i.e. the proportion of pre-reform occasionally working women who had a birth). The probability of 0.10 implies an odds ratio of $0.1/(1 - 0.1) = 0.11$

For a similar woman after 2009, the variable *time* had no effect (see Table C4), and so no adjustment has to be made for the time trend. The PLA increased the odds ratio by a factor of 1.31, meaning a post-reform odds ratio of 0.15 ($= (0.1/0.9) \cdot 1.31$). This implies that the PLA may have increased the probability of occasionally working women having a birth to 0.13 ($= 0.15/(1 + 0.15)$), a 30% increase (equivalently, a 3 percentage point increase). Similar calculations for parity 1 women imply that the PLA increased the probability of a birth from 0.32 to 0.40, a 25% increase (equivalently, an 8 percentage point increase).

5.6 Discussion

We now unpack the meaning of our results, explore the implications of our findings for policy and research, and note some limitations of our study.

Our results suggest that the PLA had no impact on the fertility of women who remained in work. This implies that the PLA did not make childbearing more attractive for women who did not intend to quit their job at any point. This finding is counter-intuitive in the sense that the PLA should facilitate childbearing for working women, by entitling them to a remunerated period of time off work at the birth of a child. Working women in Taiwan are also entitled to maternity leave. Taken together this indicates that pre-existing maternity leave entitlements are sufficient to facilitate childbearing for always working women. Working women in Taiwan may only want to take time off work at the birth of their child, and find that maternity leave is sufficient for this. Therefore the additional entitlements provided by the PLA – i.e. a longer period of leave at the birth of the child, or a period of leave taken at some other time before the child is 3 – may not be needed for working women to have children.

For occasionally working women, our results suggest that the PLA may have led to an increase in fertility, but this finding is not robust to alternative definitions of occasionally working women. We can therefore only conclude that evidence for the effect of the PLA on the fertility of occasionally working women is mixed. Any effect on occasionally working women could be because occasionally working women are less driven to work than always working women (either due to economic necessity or career ambition), and arguably they might be more eager to claim additional benefits that childbearing may entitle them to. Consequently the knowledge that an extra child would entitle someone to an additional 6 months off work at any time before that child is 3, could be an attractive proposition for such women.

Taken together, our findings suggest that expanding parental leave entitlements is not a cost-effective strategy for increasing aggregate fertility in Taiwan.

The PLA is perhaps the most expensive pro-natal policy implemented in Taiwan – although its purpose is by no means only pro-natal – and so one would hope for fertility impacts on a broader range of women. Our study is the first empirical assessment to capture the causal effect of a leave policy on fertility at the micro-level, and presents evidence which runs contrary to similar studies conducted in Europe and North America. Researchers have cautioned against understanding low fertility in East Asia in terms of historical experiences in Europe and North America, and our findings seem to reinforce this contention (Basten 2013). In terms of gender roles, our findings would seem to suggest that the fertility of highly educated women was no more responsive to the PLA than the fertility of less well educated women. This runs contrary to the argument of (Cheng and Hsu 2020), which suggests that better educated women will be more responsive to the provision of parental leave policies. Further research needs to establish what differences are important for the different fertility responses to leave in East Asia and Europe and North America – whether they are due to economic conditions, institutional arrangements or societal values.

We identify three limitations in our study due to data availability. The first limitation is that we cannot observe some covariates over the 3 year focal period that may have been important for fertility, such as income or grandparental co-residence. This is because the data is in the form of repeated cross-sections. However, the strength of our study design, as well the inclusion of some key covariates, still enables us to identify causal effects. The second limitation is that we cannot observe whether there was any self-selection of women into jobs that offered the PLA. This would be problematic for identification if women who were going to have children anyway deliberately started working to access the PLA. However, if there were such a phenomenon, it would tend to bias our findings upward, i.e. it would increase the likelihood of a (spurious) effect being measured. The fact that no effect was found indicates that even if there was self-selection of women who wanted children into jobs offering the PLA, it was insufficient to impact fertility. The third limitation is that we do not directly observe whether an individual was eligible for the PLA during the three year focal period. However, given the wide coverage of the PLA (i.e. all public and private sector workers), and the fact that we are able to observe respondent's employment sector if they are working, the assumption that respondents were eligible should not be overly restrictive.

5.7 Conclusion

The PLA has been a major and costly pro-natal policy, but there is limited evidence that it had an impact on fertility. By comparing time trends for those eligible and not eligible for the PLA, we find that the PLA had no effect on the fertility of women who were always working, but may have had some positive impact on the second births of women who were in and out of work since marriage. Our DID method – as part of a quasi-experimental design – enables the identification of causal effects, and so we can reliably claim that the PLA had no causal impact on the childbearing of always working women, and may have had a causal impact on the childbearing of occasionally working women. This study is the first to evaluate a causal impact of a national leave policy in East Asia. Contrary to the evidence from Western countries, we do not find there to be a large positive effect of leave policy on fertility. Consequently, there needs to be more research to establish whether and how parental leave increases aggregate fertility in low-fertility East Asian countries.

6

Conclusion

6.1 Introduction

The TFR in Taiwan is about 1 birth per woman, meaning that successive cohorts of childbearing-age men and women are halving in size. The persistence of lowest-low fertility for most of the past two decades mean that Taiwan's age structure now has negative momentum built into it: in 30 years there will be half the number of childbearing-age individuals that there are today, and so the population will inevitably decline to some extent. In the long run, the persistence of low fertility will have severe consequences for the Taiwanese economy and society. Rising old-age dependency ratios, higher consumption and lower investment will hamper economic growth, as will the decline in the absolute number of working-age adults. At the level of individuals and families, data on ideal family sizes – coupled with lowest-low fertility – implies that people are not having as many children as they would ideally like to have. This is problematic because failure to fulfil childbearing desires can negatively impact life satisfaction and emotional well-being.

In this thesis we had four specific research aims, relating to the effects of gender equity and leave on fertility. In this concluding chapter, we summarise the main findings of our three papers in terms of those aims. We then discuss the contributions and implications of our findings for policy and research. Next we discuss some of the limitations of our study. Finally we indicate avenues for further research

6.2 Summary of the main findings

6.2.1 The effect of housework on realised fertility in Taiwan

In Paper 1, we evaluated whether more equal divisions of housework within married couples were associated with higher probabilities of having children. We also theorised that housework division should impact fertility more in the long-term than in the short-term. This is because marital satisfaction is more responsive to housework division at younger ages, and that marital satisfaction is a key predictor of fertility. Using hybrid and classical logistic regression models on panel data, we found that there was an impact of housework balance on fertility in the long-term but not in the short-term. In the long-term models, we found the effect to be large: couples with an equal division of housework were 20.6% more likely to have another birth than couples with the mean division of housework.

6.2.2 Conceptualising the effects of pro-natal policies on fertility

Paper 2 developed a new conceptual and methodological framework, extending the analysis of (Lalive and Zweimüller 2009). The framework of Lalive and Zweimüller decomposes the effects of a parental leave policy on fertility in terms of the current-child effect, future-child effect, and total effect. We extended this framework in two ways. Firstly, we further decomposed each effect to women at different parities. Such a decomposition is conceptually necessary since parity 0 and parity 1+ women are categorically dissimilar. Secondly, we specified three study designs that are able to statistically capture each of these types of effects, and discussed which designs are at lower risk of bias. The study designs we specified are the short before-after design, the long before design, and the case-control DID design. On top of this framework, we argued that the future-child effect and total effect are of more interest to pro-natal policy-makers than the current-child effect, since the former effects concern the parental leave incentives a woman has for a child not yet born. We argued that the future-child effect and total effect are better reflections of what policy-makers mean when they refer to "the effect of a leave policy on fertility," and therefore have higher construct validity than the current-child effect.

6.2.3 The effect of parental leave policies on fertility

Having established an appropriate conceptual and methodological framework in Paper 2, we were then able to categorise and synthesise the evidence for the effect of leave on fertility. We found that the current-child effect studies mostly reported null effects, but that findings in the future-child effect and total effect studies were all supportive. In terms of effect sizes, we found that the introduction of new parental leave entitlements could increase the probability of next births by between 3-24%, with the effect size depending on the generosity of the reform. There were two limitations of our findings. Firstly, there were a comparatively small number of studies that identified the future-child effect and total effects. Secondly, almost all studies evaluated reforms in Northern and Western Europe, and Canada. This limits the generality of our findings to other settings.

6.2.4 The effect of the Parental Leave Allowance on fertility in Taiwan

In Paper 3 our focus returns to Taiwan, where we investigate whether the PLA had a causal impact on the fertility of women who were eligible for it. Following on from our conceptual framework and argument in Paper 2, we sought to statistically identify the future-child effect on women at all parities. In order to do this, we used a case-control DID study design. We evaluated two sets of models: in the first set, the 'case' group comprised of women who had always been working at least since marriage, and in the second set, the 'case' group comprised of women who had been in and out of employment since marriage. To evaluate the effect of the PLA on both sets, we compared them with the 'control' group of women who had not worked at least since marriage. In line with our findings from Paper 2, we expected those who were always working – and therefore eligible for the PLA – to have had their childbearing causally impacted by the policy. Contrary to our expectations, there was no evidence of an impact on this first set. By contrast, there was some evidence of an impact of the PLA on women who had been in and out of work since marriage, despite the fact that they were less reliably eligible for the PLA. However, our findings for the second set of women were not robust to alternative model specifications, and so cannot place too much weight on our positive findings. Overall, our findings indicate that parental leave in Taiwan may only change incentives for women who are not particularly career driven.

6.3 Contributions to policy and research

Our findings make a substantive contribution to research on low fertility, gender equity and leave policies. Here we discuss those contributions under three headings: evidence and knowledge; theory, methodology, and conceptual frameworks; and policy.

All three of our papers aim to evaluate causal impacts on realised fertility, and reliably uncover new evidence of effects or non-effects. In Paper 1, we are unable to use a strict quasi-experimental study design and therefore cannot directly test for causality. However, we are able to evaluate whether the division of housework is associated with subsequent childbearing, thereby establishing that the putative cause occurred before the putative effect. Our finding that more equitable divisions of housework are associated with higher probabilities of childbearing provides the first empirical support for gender equity theory as an explanation of low fertility in Taiwan. In Paper 2, our review of all available quasi-experimental studies demonstrates that parental leave does increase fertility in Western countries. This finding contributes to the debate on the efficacy of parental leave as a pro-natal policy, and will hopefully dispel the popular contention that the evidence is mixed. Paper 3 focusses on reliably establishing whether the PLA had a causal impact on fertility, and shows definitively that it had no effect on the fertility of women who were always working. This confounds our expectations from Paper 2, though is not inconsistent with our Paper 2 findings since Paper 3 is the first quasi-experimental evaluation of a leave policy on fertility in East Asia. We also make a contribution to the evidence of the effect of leave on fertility, by finding some indication that women who were occasionally working may have been impacted by the policy to some extent.

Papers 1 and 2 make substantive theoretical, conceptual and methodological contributions to our understanding of the effect of gender equity on fertility, and how to properly decompose and test for different types of effects of leave policies. Paper 1 posits a new theoretical argument, arguing that domestic gender inequity is likelier to have more of an impact on fertility over a longer period of time. Until now gender equity theory has been couched in static terms, and does not account for potential changes in the relationship between equity and fertility over time. By using evidence of the effect of gender equity on marital satisfaction, and of marital satisfaction on fertility, we link gender equity with fertility in such a way as to explain how this relationship might change over time. In Paper 2, we posit a conceptual and methodological framework to understand the effect of leave on fertility, specify study designs to identify each effect, and argue that certain effects are more relevant than others to pro-natal policy-makers. The purpose of this

framework is to clarify that many of the available studies of the effect of leave on fertility – while having strong study designs – are in fact only capturing a type of effect for which there is not strong theoretical reason to believe impacts fertility. The current-child effect is the effect of having leave available for a child just born, on the decision to have another child in the future. In the absence of speed premiums, the experience of having leave for a previous child seems unlikely to be a significant incentive for mothers to decide to have another child years down the line. As research interest in leave and fertility intensifies due to more countries entering low fertility regimes, these conceptual distinctions will be important for pro-natal policy evaluations. A recent quasi-experimental study – "Impact of a Reform Towards Shared Parental Leave on Continued Fertility in Norway and Sweden" (Duvander et al. 2020) – repeats the same analysis as the other current-child effect studies, and similarly finds limited evidence of an impact. With our conceptual apparatus, these researchers would have been able to understand that it would have been unlikely that they would find any effect. Lastly, our adaptation of ROBINS-I for evaluating the idiosyncrasies of pro-natal policy evaluations could be useful for systematic reviews of other pro-natal policies, and maybe reviews of other types of social policy.

In policy terms Papers 2 and 3 both concern the evaluation of parental leave policies, so insofar as those papers have generalisable results, they should contribute to analyses of parental leave policies. For other low fertility countries in East Asia, our results in Paper 3 indicate that leave entitlements of 6 months are not sufficient to impact fertility. For Taiwan specifically, Paper 3 indicates that other pro-natal policies may be more cost-effective. For Western countries, our findings in Paper 2 indicate that parental leave is a reliable method for increasing fertility.

6.4 Limitations

The research conducted over the course of the PhD was limited in a few ways, mostly relating to data availability and sample surveys not asking questions about certain outcomes of interest. Here we discuss some limitations imposed by data availability and the exclusion of unmarried women from the study, consider specific limitations of each of the three papers, and conclude with a discussion of the limitations imposed by language barriers.

In terms of data availability, there was a trade-off between survey sample size and survey detail. On the one hand, the Women's Marriage, Fertility and Employment Surveys (WMFES) provided very large cross-sectional data, with about 20,000 respondents in each wave. However, there are no detailed ques-

tions on key variables relating to gender equity theory – such as housework and childcare – in the WMFES. In contrast, the Panel Survey of Family Dynamics (PSFD) and the Taiwan Social Change Survey (TSCS) ask much more detailed questions about domestic labour and childbearing behaviours, but have much smaller sample sizes. Unlike many other high-income countries, Taiwan lacks a large, longitudinal dataset covering family behaviours. Consequently this limited our ability to investigate a range of different factors relating the effects of domestic and formal labour on childbearing, and the small size of the PSFD made subgroup analysis difficult. In terms of administrative data, we did not have access to administrative datasets on eligibility for the PLA or on fertility, meaning that eligibility for the PLA had to be estimated.

The analysis throughout Papers 1 and 3 were restricted to married heterosexual couples, and therefore represents a bias in our results towards married women. Consequently, this means that we cannot shed light on the fertility behaviour of unmarried women, divorced women, or same-sex couples. To a large extent this restriction was unavoidable, since the WMFES did not ask unmarried women about their childbearing behaviour until 2016. We were therefore not able to use the WMFES to study the fertility of unmarried women. Additionally, restricting our analysis to only married heterosexual couples was motivated by the fact that much of gender equity theory is specified for heterosexual couples. Gender equity theory is largely silent on the intergenerational balance of housework between genders, and its impact (if any) on fertility. Consequently, our scope was limited by data availability, theory, and the fact that marriage remains a *de facto* prerequisite for childbearing in Taiwan.

In Paper 1, we investigate the impact of domestic gender equity on fertility. Ideally, it would have been possible to evaluate a range of different domestic labour tasks, and specifically both childcare and housework. Unfortunately the PSFD does not ask the required questions on childcare, meaning that we were limited to using housework alone. We were further limited by the fact that the PSFD asks individuals to retrospectively report the hours of housework done that week by them and their spouse. This is not an ideal measurement of housework, as time-use diaries have shown to be more accurate. Another central limitation was imposed by the relatively small size of the PSFD dataset, meaning that it was difficult to investigate any important subgroup differences. Furthermore, this meant there were likely too few births from 2012-2015 to establish any impact of domestic labour balance on fertility over that period. Lastly, we were limited by the fact that the coverage of the data was only over a five year period, meaning that we were limited in the extent to which we could investigate how housework sharing might impact fertility differently over different time scales.

In Paper 2, we investigate the impact of parental leave policies on fertility. Our studies were drawn entirely from Europe and North America, meaning that we could not reasonably generalise those findings to Taiwan or other East Asian countries. Essentially, this limitation was imposed by there being no existing quasi-experimental evaluations of pro-natal policies in the English-language literature. This could partly be due to the shorter time frames that East Asian countries have had post-transitional fertility rates. Nonetheless, the paper is still useful for studies evaluating the impact of policies on fertility, due to the conceptual framework it develops.

Paper 3 evaluates the impact of the Parental Leave Allowance (PLA) on fertility. We use the WMFES data to evaluate this effect. Though we were able to estimate respondents' career histories, the fact that the WMFES is a cross-sectional dataset meant that those histories could not be directly observed. Consequently, there will inevitably have been some error in determining precise employment histories. Beyond employment histories, eligibility for the PLA was also not directly observed, and had to be estimated by evaluating current and previous employment sectors. As mentioned previously, if we had had access to PLA registration data – as well as administrative vital data – then this would have improved the study.

Beyond limitations in terms of data and scope, I was also hindered by language. I cannot speak Chinese, and so that may have limited my ability to engage with literature written in Chinese, particularly research published in Chinese-language journals. However, many Chinese-language journals provide titles or abstracts in English, and so I was able to browse that literature for articles that may have been potentially relevant to the focus of my thesis. For legal documents on pro-natal policies and documentation for surveys used in the research, a combination of Google Translate and asking colleagues for assistance was sufficient to cover most issues.

6.5 Further research

Our findings indicate several avenues for future research. The conceptual framework established in Paper 2 could be applied to systematic reviews of other pro-natal policies such as baby bonuses and childcare, which are only made available to parents after the birth of a child. Our framework could also be used for framing individual studies of parental leave in other settings, providing a rubric for which effects are being identified, what study design is being adopted to identify those effects, and whether the common sources of RoB have been controlled for or mitigated. Our finding in Paper 3 – that the PLA did not impact fertility – is the

first quasi-experimental study of a national parental leave policy on fertility in East Asia. It therefore needs to be established whether our findings generalise to other East Asian countries, and if so, why there is a systematic difference in the effect of leave policies on fertility in East Asia, and in Europe and North America. Lastly, our findings in Paper 3 indicate that the PLA may have impacted the fertility of occasionally working women but not always working women. This needs to be theoretically explained, since both sets of women face the same types of incentives.

Appendices

A

Paper 1 Appendices

A.1 Full hybrid model results

(Table on next page)

Table A1: Full range of hybrid panel data logistic regression estimates of births and desire for more births, by survey respondent group.

		Births			Desire for more births		
		(1)	(2)	(3)	(4)	(5)	(6)
Type of estimate	**Variable**	**All**	**Females**	**Males**	**All**	**Females**	**Males**
Time-invariant variable estimates	Gender	-0.336			-0.174		
		(0.289)			(0.258)		
	Wife's tertiary education		-0.175			0.151	
			(0.705)			(0.468)	
Within estimate	Youngest child younger than five	-1.108*	-1.268	-0.999	-0.966**	-1.408**	-0.514
		(0.559)	(0.930)	(0.805)	(0.365)	(0.525)	(0.530)
	Wife's tertiary education	0.360		1.399+	0.150		0.204
		(0.630)		(0.775)	(0.405)		(0.493)
	Husband's tertiary education	1.280*	1.663+	-5.143*	-0.0419	-0.214	1.194
		(0.630)	(0.913)	(2.358)	(0.416)	(0.504)	(1.405)
	Proportion of housework done by husband	0.983	1.365	-0.160	-0.685	-1.281	0.352
		(1.199)	(1.806)	(1.970)	(0.833)	(1.159)	(1.264)
	Wife's hours of housework	-0.0183	-0.0353	-0.0137	0.00560	0.0163	0.00116
		(0.0256)	(0.0430)	(0.0390)	(0.0146)	(0.0230)	(0.0196)
	Husband's hours of housework	-0.0462	-0.0754	-0.00756	0.0307	0.0533+	-0.0000332
		(0.0333)	(0.0517)	(0.0523)	(0.0206)	(0.0279)	(0.0319)
	Wife's age	0.621***	0.558*	0.951***	-0.616***	-0.683***	-0.576***
		(0.153)	(0.239)	(0.247)	(0.0949)	(0.142)	(0.137)
	Wife works	0.467	0.133	1.284	1.040*	1.685*	0.538
		(0.687)	(1.251)	(0.913)	(0.442)	(0.702)	(0.589)
	Parents co-reside	0.744	-0.238	1.950*	-0.134	0.181	-0.412
		(0.599)	(0.993)	(0.886)	(0.389)	(0.536)	(0.585)
	Log income	0.137	0.129	0.239	0.0169	0.0121	-0.00472
		(0.108)	(0.141)	(0.315)	(0.0597)	(0.0727)	(0.118)
	Parity 2 or more	-7.841***	-8.555***	-8.866***	-0.502	-0.372	-0.793+
		(0.707)	(1.165)	(1.116)	(0.327)	(0.492)	(0.454)
	Pays for nanny	0.758	-0.599	1.646*	-0.250	0.809	-1.056*
		(0.538)	(0.987)	(0.772)	(0.372)	(0.577)	(0.516)
	Year of the Dragon	0.542+	0.755	0.595			
		(0.292)	(0.483)	(0.410)			
	Desires more children	-0.509	0.227	-1.669*			
		(0.511)	(0.855)	(0.763)			
	Parents co-reside * pays for nanny	-1.540+	-1.194	-1.689	-0.108	-0.737	0.270
		(0.844)	(1.468)	(1.261)	(0.556)	(0.830)	(0.766)
	Wife works * wife's hours of housework	-0.00251	0.0373	-0.0699+	-0.0319+	-0.0564*	-0.00949
		(0.0245)	(0.0407)	(0.0388)	(0.0164)	(0.0250)	(0.0233)
	Desires more children * wife's hours of housework	0.0734**	0.101**	0.0818*			
		(0.0231)	(0.0381)	(0.0358)			
Between estimate	Youngest child younger than five	3.549***	2.984**	4.145***	4.154***	3.340***	5.016***
		(0.771)	(1.126)	(1.201)	(0.458)	(0.615)	(0.722)
	Wife's tertiary education	-0.0434		0.874	-0.175		-0.592
		(0.410)		(0.693)	(0.358)		(0.598)
	Husband's tertiary education	0.554	-0.162	0.804+	0.230	0.563	0.0183
		(0.377)	(0.931)	(0.476)	(0.340)	(0.605)	(0.436)
	Proportion of housework done by husband	0.414	0.170	0.557	-0.110	1.502	-2.694
		(1.523)	(2.424)	(2.491)	(1.377)	(1.740)	(2.290)
	Wife's hours of housework	-0.00325	0.0269	-0.0526	-0.00138	0.0289	-0.0137
		(0.0280)	(0.0427)	(0.0443)	(0.0234)	(0.0344)	(0.0332)
	Husband's hours of housework	-0.0582	-0.0487	-0.0703	0.0139	-0.0165	0.0909
		(0.0437)	(0.0695)	(0.0740)	(0.0389)	(0.0506)	(0.0645)
	Wife's age	-0.0540	-0.0829	-0.0618	-0.135***	-0.219**	-0.119*
		(0.0456)	(0.0973)	(0.0592)	(0.0403)	(0.0746)	(0.0500)
	Wife works	-0.508	-0.702	-0.837	0.228	-0.300	1.311
		(0.737)	(1.215)	(1.214)	(0.676)	(0.876)	(1.096)
	Parents co-reside	0.497	0.808	0.208	-0.378	-0.218	-0.363
		(0.370)	(0.684)	(0.524)	(0.320)	(0.464)	(0.459)
	Log income	-0.133	-0.119	0.134	0.0368	-0.0145	0.315
		(0.139)	(0.184)	(0.346)	(0.120)	(0.135)	(0.260)
	Parity 2 or more	0.165	0.0736	0.406	-3.012***	-2.874***	-3.391***
		(0.436)	(0.712)	(0.646)	(0.323)	(0.437)	(0.497)
	Pays for nanny	0.841	3.042**	-0.480	-0.868	-1.086	-0.975
		(0.597)	(0.986)	(0.862)	(0.575)	(0.904)	(0.783)
	Year of the Dragon	0	0	0			
		(.)	(.)	(.)			
	Desires more children	0.545	1.019	-0.387			
		(0.786)	(1.261)	(1.159)			
	Parents co-reside * pays for nanny	-0.325	-1.870	0.470	1.242	1.218	1.495
		(0.854)	(1.462)	(1.213)	(0.865)	(1.280)	(1.222)
	Wife works * wife' s hours of housework	0.0283	-0.00368	0.0582	0.00530	-0.0144	-0.0176
		(0.0337)	(0.0542)	(0.0559)	(0.0308)	(0.0422)	(0.0488)
	Desires more children * wife's hours of housework	-0.0237	-0.0371	0.0107			
		(0.0375)	(0.0563)	(0.0598)			
	Constant	-3.308	-2.937	-6.995+	1.141	4.250	-3.044
		(2.199)	(3.853)	(3.987)	(1.846)	(2.827)	(3.063)
	N	2176	1040	1136	2176	1040	1136

Notes: Within estimates are derived using variation within individuals over time, and between estimates use variation across different individuals. Estimates of the log-odds, standard errors (in brackets), and odds ratios are reported to three decimal places. The odds ratios are the exponents of the log-odds. $+p < .10$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

A.2 Full cross-sectional model results

(Table on next page)

Table A2: Full range of cross-sectional logistic regression estimates of births, by survey respondent group.

Variable	(7) **All**	(8) **Females**	(9) **Males**
Youngest child younger than five	0.631+	0.917	0.311
	(0.377)	(0.574)	(0.546)
Wife's tertiary education	0.854*	0.518	1.180*
	(0.366)	(0.536)	(0.543)
Husband's tertiary education	0.156	0.363	0.0881
	(0.333)	(0.484)	(0.519)
Proportion of housework done by husband	1.051	3.560**	-2.058
	(1.040)	(1.246)	(1.556)
Wife's hours of housework	-0.0546*	-0.0452	-0.0784**
	(0.0228)	(0.0467)	(0.0268)
Husband's hours of housework	-0.00846	-0.0601+	0.0664+
	(0.0266)	(0.0309)	(0.0398)
Wife's age	-0.286***	-0.397***	-0.270***
	(0.0537)	(0.0971)	(0.0653)
Wife works	-1.381*	-2.002*	-0.674
	(0.561)	(0.963)	(0.681)
Parents co-reside	-0.0291	0.0695	-0.113
	(0.302)	(0.443)	(0.423)
Log income	0.0648	0.128	-0.0149
	(0.0672)	(0.0834)	(0.108)
Parity 2 or more	-2.202***	-2.035***	-2.563***
	(0.294)	(0.421)	(0.464)
Pays for nanny	0.573	0.346	0.938
	(0.419)	(0.632)	(0.642)
Desires more children	0.627	0.252	0.979+
	(0.400)	(0.621)	(0.580)
Parents co-reside * pays for nanny	-0.220	-0.278	-0.240
	(0.615)	(0.831)	(1.018)
Wife works * wife's hours of housework	0.0596*	0.0788	0.0226
	(0.0264)	(0.0511)	(0.0305)
Desires more children * wife's hours of housework	0.0228	0.0303	0.0255
	(0.0187)	(0.0299)	(0.0246)
Gender	-0.0793		
	(0.263)		
Constant	7.991***	10.29**	8.810***
	(1.887)	(3.225)	(2.471)
N	544	260	284

Notes: The dependent variable here is whether the couple went on to have one or more children in 2011-2015. Estimates and standard errors (in brackets) are reported to three decimal places. $+p < .10$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

B

Paper 2 Appendices

B.1 PROSPERO systematic review protocol

The review protocol is available online at PROSPERO, at https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=128493. The protocol was registered on 09/09/2019. We reproduce the printed version here, starting on the next page.

The effect of parental leave policies on increasing fertility: a systematic review

Jac Thomas, Francisco Rowe, Eric Lin

Citation

Jac Thomas, Francisco Rowe, Eric Lin. The effect of parental leave policies on increasing fertility: a systematic review. PROSPERO 2019 CRD42019128493 Available from: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42019128493

Review question

What does the evidence show for how changes in leave entitlements affect fertility? By 'leave,' we mean parental leave, maternity leave, and paternity leave. Do changes in leave entitlements affect fertility quantum, fertility tempo, or both? How do different types of changes in leave entitlement affect fertility? How do heterogeneous incentive changes affect fertility in different groups?

Searches

The following bibliographic databases will be used to search for relevant studies and reviews. Since the breadth of eligible countries and time periods is large, we will not restrict the search on the basis of time periods and countries; this filtering will be done later, if needed. Our search is restricted to material available in English. The full search strategy is linked to below.

Academic sources:

PopLine
Academic Search Premier (multi-disciplinary)
ProQuest dissertation & theses A&I
PubCentral (medical, full-text)
PubMed (medical science)
Science Citation Index (science and technology journals)
Social Science Citation Index (social science journals)
The Cochrane Library (Cochrane Database for Systematic Reviews, other reviews with a medical focus)
Social Care Online (UK database, social science)
Business Source Elite
EconLit
IDEAS
Economist Online
Campbell Collaboration Library
Centre for Reviews and Dissemination Databases

EPPI-Centre Systematic Reviews

Grey literature:

OpenGrey (http://www.opengrey.eu/)
Social Science Research Network (https://www.ssrn.com/en/)
WorldWideScience (https://worldwidescience.org/)

International Network on Leave Policies & Research (https://www.leavenetwork.org/)
UN World Population Policies Database (https://esa.un.org/poppolicy/about_database.aspx)
Office of Economic Coordination and Development (OECD) (http://www.oecd.org)

Economic and Social Research Council (ESRC) (https://esrc.ukri.org/)
Centre for Economic Policy and Research (CEPR) (https://cepr.org/)

Types of study to be included
All studies that aim to quantify the effect of a parental leave policy change, using micro-level data.

Condition or domain being studied
We are interested in fertility: the number of births had by women. We are interested in how changes in leave entitlements affect the number of births had by women.

Participants/population
All women of childbearing age. This includes women who are married, cohabiting, or neither married nor cohabiting.

Intervention(s), exposure(s)

Parental, maternity, and paternity leave: time off from work given to parents, mothers, and fathers, respectively, under the condition that the individual taking the leave can return to their job after the period of leave is completed. It may be paid or unpaid.

Comparator(s)/control
At the moment, we are open to all empirical strategies that aim to evaluate the effect of a leave policy change on fertility.

Main outcome(s)
Fertility - the number of children had by women. We are interested in both tempo and quantum effects. We are interested at fertility at any time period after the leave policy change.

* Measures of effect

The effect is measured by whether women have more children. Birth tempo will also be considered.

Additional outcome(s)
None.

* Measures of effect

Not applicable.

Data extraction (selection and coding)

1. Search results from the academic databases and other sources will be combined, and duplicates will be removed. Each duplicate found by the reference management software will be considered manually – we won't automatically remove duplicates.
2. Screen the search results based on titles and abstracts. For each included article, the title will be read – if it's clearly irrelevant, it will be excluded. If it may be relevant, the abstract will be read. If the abstract shows it to be clearly irrelevant, it will be excluded. If the abstract makes shows the article to be clearly relevant, or if it's not clear whether it's relevant or irrelevant, the article will be included for the next stage.
3. The full text of each article included will be retrieved.
4. Screen those articles. The reasons for excluding any seemingly relevant articles will be specified in an appendix to the final report.
5. The resulting articles will then be coded, according to the coding form below. The coding form will be adjusted if deemed necessary.

Data to be extracted:

Names of author(s)

Title
Language
Journal
Year
Country
Participant characteristics (age, gender, parity, employment, marital status, socioeconomic criteria, region, ethnicity)
Time period covered (date of policy change, length of follow-up)
Details of initial policy (payment, time length, timing and other incentives)
Details of new policy (payment, time length, timing and other incentives)
Statistical difference generated (e.g. case-control before-after, case-control regional, heterogeneous incentives)
Type of data used (e.g. panel, cross-section, repeated cross-section)
Source of data (e.g. administrative, questionnaire)
Level of aggregation (e.g. mother, family)
Sample size (treated, comparison)
Key dependent variables
Key independent variables

Risk of bias (quality) assessment

In order to assess the risk of bias in studies included in the review, we will use the ROBINS-I tool for the assessment of risk of bias in non-randomised studies (Sterne et al. 2016). ROBINS-I builds on the Cochrane Collaboration's guidelines for assessing the risk of bias in randomised studies, and specifies seven domains of possible bias. A series of questions are provided for each of the seven domains, to be asked of every study included in the review. Based on the responses to these questions, each study is then given an overall score for risk of bias.

A single researcher (Jac Thomas) will conduct the risk of bias analysis. Since no quantitative synthesis of the studies will be conducted, assessment of risk of bias for included studies is not expected to influence the synthesis of the studies.

Strategy for data synthesis

A narrative (descriptive) synthesis will be conducted. A quantitative synthesis or meta-analysis won't be possible, due to heterogeneity in study designs, populations and interventions. Individual-level participant data from included studies will not be sought out or used. Only aggregate data presented in the published form of included studies will be included.

The narrative synthesis will be conducted using guidance summarised in:

Ryan R; Cochrane Consumers and Communication Review Group. 'Cochrane Consumers and Communication Review Group: data synthesis and analysis'. http://cccrg.cochrane.org, 13 March 2019 (accessed 03/09/2019).

Specifically, we will synthesise the results using a four step approach:

1. Developing a theory about how the intervention works, why and for whom (using included studies)

2. Conducting an initial synthesis that summarises the results of included studies, using information extracted from coding. This will include commenting on any differences in the results of studies using different designs, and using vote counting

3. Discussing how differences of effects found within and between studies may be explained by study heterogeneity

4. Assessing the robustness of the synthesis, given the amount and quality of the evidence, and the methods used for synthesising the evidence.

Analysis of subgroups or subsets
No separate analysis of subgroups is planned.

Contact details for further information
Jac Thomas
jathomas@liverpool.ac.uk

Organisational affiliation of the review
University of Liverpool
https://www.liverpool.ac.uk/

Review team members and their organisational affiliations
Mr Jac Thomas. University of Liverpool
Dr Francisco Rowe. University of Liverpool
Professor Eric Lin. National Tsing-Hua University, Taiwan

Collaborators
Ms Lisa Hawksworth. University of Liverpool

Type and method of review
Intervention, Narrative synthesis, Systematic review

Anticipated or actual start date
18 March 2019

Anticipated completion date
31 July 2019

Funding sources/sponsors
University of Liverpool; National Tsing-Hua University, Taiwan

Conflicts of interest

Language
English

Country
England, Taiwan

Stage of review
Review Ongoing

Subject index terms status
Subject indexing assigned by CRD

Subject index terms
Fertility; Humans; Parental Leave; Policy

Date of registration in PROSPERO
09 September 2019

Date of first submission
03 June 2019

Stage of review at time of this submission

Stage	Started	Completed
Preliminary searches	Yes	No
Piloting of the study selection process	No	No
Formal screening of search results against eligibility criteria	No	No
Data extraction	No	No
Risk of bias (quality) assessment	No	No
Data analysis	No	No

The record owner confirms that the information they have supplied for this submission is accurate and complete and they understand that deliberate provision of inaccurate information or omission of data may be construed as scientific misconduct.

The record owner confirms that they will update the status of the review when it is completed and will add publication details in due course.

Versions
09 September 2019

B.2 Deviations from the review protocol

Academic database searches

- POPLINE was not searched because it no longer exists
- Social Science Research Network (SSRN) was not searched because it does not support the use of wildcards or Booleans in searching
- Social Care Online was not searched because it does not support advanced searching
- The study design terms were removed because they led to articles we knew to be relevant being omitted
- Some of the search terms were removed from the fertility group, as they were deemed unnecessary

Grey literature searches

- Social Science Research Network (SSRN) was not searched because it does not support the use of wildcards or Booleans in searching
- WorldWideScience was not searched, because it cannot limit searches based on title and/or abstract
- Only the first 200 results from the Google Scholar search were looked through, because the search returned thousands of results
- The INLPR country reports were retrieved, and the "recent research" sections were browsed for any mention of empirical articles investigating the effect of leave on fertility
- Additional UN sources were searched, including the Department of Economic and Social Affairs, the World Health Organisation, and the United Nations Development Programme

B.3 Search strategy for Academic Search Complete

Searched Academic Search Complete through EBSCOhost. Searched on 12/10/2019, returned 376 results.

Language: English. Search mode: Boolean/Phrase. After search, Limited by Type: Academic Journals (omitting Magazines, Newspaper articles, and Trade publications).

Title OR Subject Terms OR Abstract or Author-Supplied Abstract OR Author-Supplied Keywords.

((fertility OR birth* OR child#birth OR child#bearing OR natal*) AND ("parental leave" OR "maternity leave" OR "paternity leave" OR "family policy" OR "family policies"))

B.4 Domains of bias in ROBINS-I

1. *Confounding* occurs when a variable that predicts the outcome also predicts the intervention.

2. *Selection bias* refers to bias arising from exclusion of eligible participants, when the exclusion criteria are related to both the intervention and the outcome.

3. *Misclassification bias* occurs when study participants are erroneously recorded as having received or not received the intervention.

4. *Performance bias* refers to bias arising from deviations from intended interventions between the treatment and control group. If there are important co-interventions given to either the treatment or control group – such as the simultaneous introduction of another pro-natal policy – then the study will be at risk of performance bias.

5. *Bias due to missing data* could arise due to survey participant attrition, or from key variables (such as employment status) not being directly observed.

6. *Detection bias* concerns whether study participants are aware of whether or not they receive the intervention. Detection bias is not applicable to studies of public policy effects, since all participants are aware of whether they are eligible for a policy or not.

7. *Outcome reporting bias* could arise if only the results from models that have significant results are reported.

C

Paper 3 Appendices

C.1 Employment status in the focal period

Table 5.1 presents information on the 10 employment history categories in the WMFES. For Groups (1), (4), and (10), we can immediately determine whether respondents were working or not working during the focal period. However, it is not immediately clear whether respondents in the other groups were working during the focal period. For most of these other groups, we can use data from other questions in the PLA to estimate whether respondents were likely to be working or not working. In what follows, we discuss why we classified each of these other groups as "likely working," "likely not working," or "unclear." Our analysis here uses data from the WMFES 2016 questionnaire only.

Group (2): Quit work at marriage and have a job now

For Group (2) women, we can use the variable *b1_a* to find out when they got married, and variable *b9a* to find out how long before they returned to work. Using this information, we find that 88% of Group (2) women had returned to work before the focal period, and 95% had returned to work by the end of the focal period. However, these women may have quit at some point after returning to work after marriage, and so we cannot be certain that they were within work during the focal period. Nevertheless, the high percentages of 88% and 95% suggest that the majority of them will have been in work during the focal period, and so we classify Group (2) women as being "likely working" during the focal period.

Group (3): Quit work at marriage, had a job after but don't have a job now

The analysis for Group (3) women is similar to the Group (2) women, i.e. using information on their year of marriage and length of time before returning to work to calculate what percentage of women had returned to work at least once before the focal period. We calculate that 90% of Group (3) women had returned to work at least once before the focal period, and that 100% of Group (3) women had returned to work at least once by the end of the focal period. However, we do know that all women in Group (3) were out of work in 2016, meaning that some of them were likely to be out of work during the focal period. We therefore classify Group (3) women as being "likely working" during the focal period.

Group (5): Quit work at birth and have a job now

We estimate that parity 0 women in Group (5) were working in the focal period. We know that these women must have had at least one birth by the survey year, and will have quit work at one of those births, so prior to that they would have been working. Typically then, a parity 0 Group (5) woman would have been working at the start of the focal period (because they did not quit work at marriage), would have had a child, and then quit their job. For parity 1+ Group (5) women, we can use data from variable *b13a* to find out which birth they first quit their job at (i.e. first birth, second birth etc.), and data from variables *b2a1_a* to *b2a1_d* to find out the dates of birth of their children. We can then use variable *b12a* to find out how long they were out of work after the first time they quit due to a birth, and thereby determine the year in which they returned to work. From this analysis, we can see that 79% of Group (5) had returned to work at least once before the focal period, and 87% had returned to work at least once by the end of the focal period. We therefore classify Group (5) women as being "likely working" during the focal period.

Group (6): Quit work at birth

Group (6) women quit work at birth and were not working in 2016, though they might have re-entered employment between quitting at birth and 2016. We estimate that parity 0 women in Group (6) were working in the focal period, for the same reason that parity 0 Group (5) women were working in the focal period. For women at parities 1-3, we can cross-tabulate parity with variable *b10*, which specifies whether they ever returned to work after quitting at birth. 91% of parity 1 women never returned to work, meaning that they wouldn't have been working

during the focal period. The percentages for parity 2 and parity 3 women are 84% and 85% respectively. We can also use data from variable *b13a*, which asks respondents at which birth did they quit birth. This reveals that 87% of parity 1+ respondents quit work at their first birth. Together, these percentages indicate that the vast majority (90%) of parity 1+ Group (6) women typically quit work at their first birth and never returned to work, meaning that we can classify parity 1+ Group (6) women as likely not working during the focal period.

Group (7): Quit work for other reasons and have a job now

Unlike for women who quit work due to marriage or childbearing, there are no questions of the form, 'how long did it take for you to become re-employed after quitting?" Data from variable *b15c* – the main reason for quitting – shows a wide variety of reasons, like childcare, low pay, and redundancy. Most of these women returned to work to help the family finances or to get an independent income (*b15c1_a*. We therefore cannot estimate whether Group (7) were working in the focal period, and so we classify their focal period employment status as "unclear."

Group (8): Quit work for other reasons

Similar to Group (7) women, there is no data on how long Group (8) women were out of work (or if they ever returned to work). Therefore it is not possible to estimate whether they were in work or not during the focal period. We therefore classify the focal period employment status of Group (8) women as "unclear."

Group (9): No job before marriage, but had a job after marriage until now

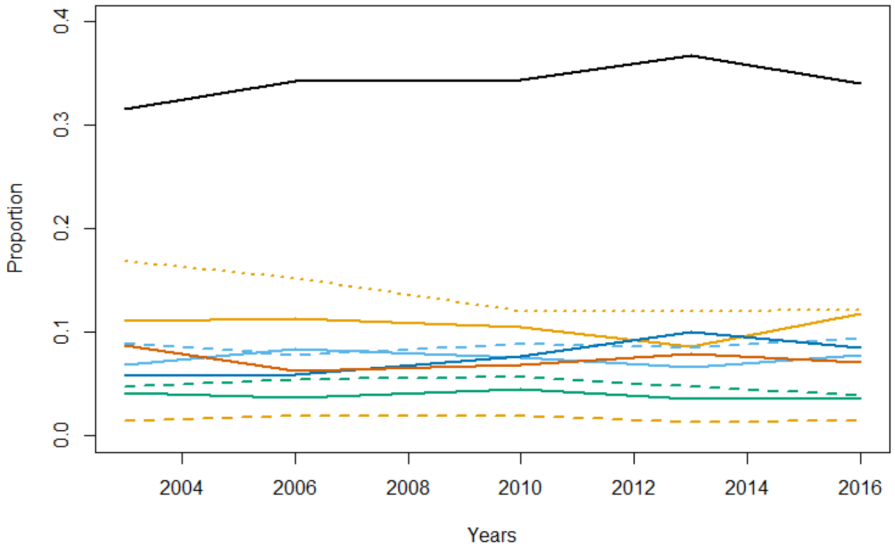
For women in Group (9), what we need to know is how soon these women entered work after marriage – specifically, whether they were working while they were having children, and whether they were working during the focal period. It could be the case that these women – who had never worked before marriage – got married, then remained out of work for a few years while they had children, and then returned to work at the end. Alternatively, it could be the case that these women got married, immediately started working, and then had children while they were working. Unfortunately, there are not any questions that could shed light on whether either of these situations were true. Consequently, it is not possible for us to evaluate whether these women were likely to be working or not working during the focal period. This means that this group should be treated with caution – we just cannot know whether they were working or not during the

focal period. However, what we do know is that they are working at the time of the survey. We therefore classify the focal period employment status of these women as "unclear."

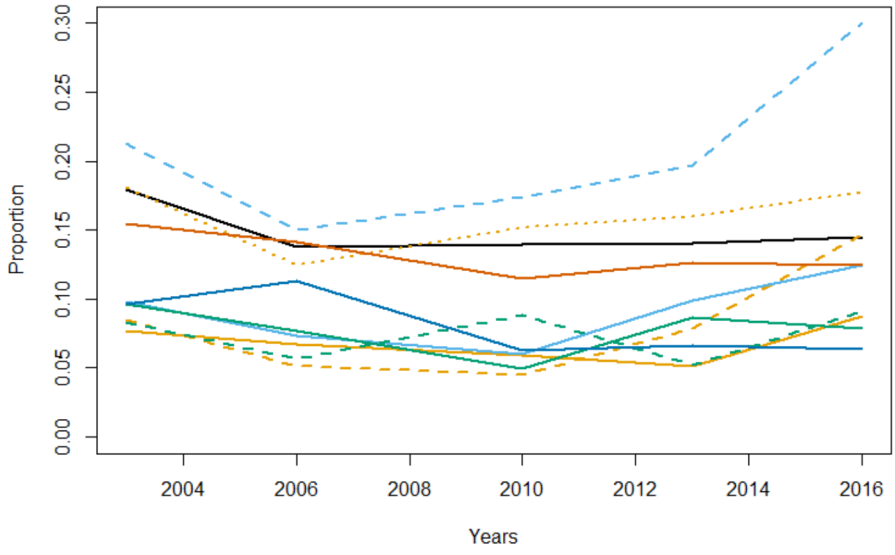
C.2 Trends in all employment history groups

(Figure on next page)

(a) Proportion of women in each employment history category



(b) Proportion of women having a birth in the 3 year focal period, by employment history



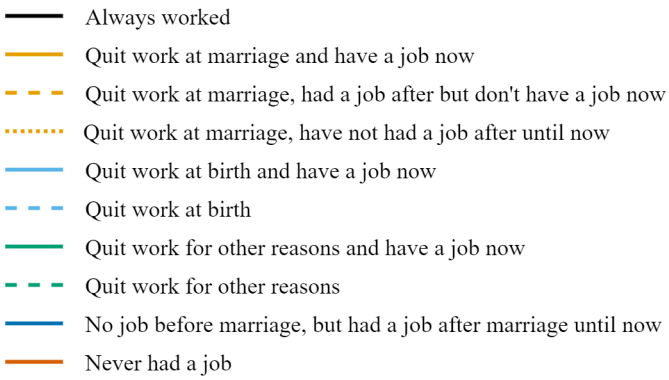


Figure C1: Proportions and fertility of wives by employment history.

Notes: The numbers of respondents in each of the five surveys from 2003 to 2016 are as follows: 9329, 8199, 7234, 6146, and 5357.

C.3 Full model results

Table C1: Always working women (Group (A)) – models without covariates.

	All women	Parity 0	Parity 1	Parity 2	Parity 3
(Intercept)	-1.65	0.16	-0.28^{***}	-2.54^{***}	-3.83^{***}
	(0.89)	(0.11)	(0.08)	(0.11)	(0.18)
Time	-0.40^{***}	-0.47^{**}	-0.44^{***}	-0.19	-0.01
	(0.08)	(0.17)	(0.12)	(0.16)	(0.38)
Treat	-0.30^{***}	-0.01	-0.26^{**}	-0.60^{***}	-0.85^{**}
	(0.06)	(0.12)	(0.08)	(0.11)	(0.29)
DID	-0.05	-0.20	0.00	-0.04	-1.20
	(0.10)	(0.20)	(0.14)	(0.21)	(0.84)
Parity effects					
– Parity 0	2.06				
	(0.96)				
– Parity 1	1.47				
	(0.95)				
– Parity 2	-0.93				
	(0.95)				
– Parity 3	-2.32				
	(0.96)				
N	21,741	2,236	4,756	10,060	4,689

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$.

Table C2: Always working women (Group (A)) – models with covariates.

	All women	Parity 0	Parity 1	Parity 2	Parity 3
(Intercept)	5.25^{***}	7.05^{***}	5.91^{***}	4.74^{***}	4.27^{***}
	(0.71)	(0.41)	(0.26)	(0.34)	(0.82)
Time	-0.03	0.01	-0.10	0.12	0.38
	(0.09)	(0.20)	(0.14)	(0.17)	(0.40)
Treat	-0.32^{***}	-0.10	-0.31^{**}	-0.47^{***}	-0.72^{*}
	(0.07)	(0.15)	(0.10)	(0.12)	(0.31)
DID	-0.15	-0.30	-0.06	-0.12	-1.46
	(0.11)	(0.24)	(0.16)	(0.22)	(0.85)
Wife's age	-0.21^{***}	-0.23^{***}	-0.21^{***}	-0.21^{***}	-0.22^{***}
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)
Wife's education: senior high school	0.57^{***}	0.82^{***}	0.65^{***}	0.32^{*}	0.29
	(0.08)	(0.17)	(0.11)	(0.14)	(0.31)
Wife's education: university	0.90^{***}	1.06^{***}	1.03^{***}	0.57^{**}	0.44
	(0.10)	(0.19)	(0.14)	(0.21)	(0.69)
Husband's education: senior high school	0.24^{***}	0.30^{*}	0.32^{***}	0.11	-0.18
	(0.07)	(0.13)	(0.10)	(0.13)	(0.31)
Husband's education: university	0.49^{***}	0.26	0.64^{***}	0.46^{*}	0.58
	(0.08)	(0.15)	(0.12)	(0.19)	(0.61)
Parity effects					
– Parity 0	1.56				
	(0.67)				
– Parity 1	1.09				
	(0.67)				
– Parity 2	-0.90				
	(0.67)				
– Parity 3	-1.69				
	(0.70)				
N	21,741	2,236	4,756	10,060	4,689

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table C3: Occasionally working women (Group (B)) – models without covariates.

	All women	Parity 0	Parity 1	Parity 2	Parity 3
(Intercept)	-1.60	0.13	-0.28^{**}	-2.56^{***}	-3.85^{***}
	(0.86)	(0.10)	(0.09)	(0.11)	(0.19)
Time	-0.40^{***}	-0.46^{**}	-0.44^{***}	-0.17	-0.00
	(0.08)	(0.17)	(0.12)	(0.16)	(0.38)
Treat	-0.34^{***}	0.27	-0.41^{***}	-0.57^{***}	-0.71^{**}
	(0.06)	(0.14)	(0.09)	(0.11)	(0.26)
Did	0.35^{***}	0.24	0.41^{**}	0.03	0.72
	(0.11)	(0.23)	(0.15)	(0.21)	(0.49)
Parity effects					
– Parity 0	2.06				
	(0.82)				
– Parity 1	1.29				
	(0.81)				
– Parity 2	-1.07				
	(0.81)				
– Parity 3	-2.25				
	(0.82)				
N	$20,018$	$1,339$	$3,816$	$9,503$	$5,360$

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table C4: Occasionally working women (Group (B)) – models with covariates.

	All women	Parity 0	Parity 1	Parity 2	Parity 3
(Intercept)	5.73^{***}	7.44^{***}	6.29^{***}	4.94^{***}	5.27^{***}
	(0.68)	(0.52)	(0.29)	(0.34)	(0.74)
Time	0.01	0.02	−0.09	0.16	0.48
	(0.09)	(0.21)	(0.14)	(0.17)	(0.41)
Treat	-0.20^{**}	0.21	-0.32^{**}	-0.27^{*}	−0.36
	(0.07)	(0.17)	(0.10)	(0.12)	(0.28)
DID	0.28^{*}	0.27	0.36^{*}	−0.02	0.64
	(0.12)	(0.28)	(0.17)	(0.22)	(0.51)
Wife's age	-0.22^{***}	-0.23^{***}	-0.21^{***}	-0.22^{***}	-0.24^{***}
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)
Wife's education: senior high school	0.44^{***}	0.61^{**}	0.56^{***}	0.24	0.41
	(0.08)	(0.20)	(0.12)	(0.13)	(0.27)
Wife's education: university	0.86^{***}	0.87^{***}	1.11^{***}	0.59^{*}	−0.31
	(0.11)	(0.25)	(0.16)	(0.23)	(0.71)
Husband's education: senior high school	0.11	0.14	0.27^{**}	−0.05	−0.44
	(0.07)	(0.17)	(0.10)	(0.12)	(0.26)
Husband's education: university	0.39^{***}	0.27	0.56^{***}	0.10	0.79
	(0.10)	(0.22)	(0.14)	(0.20)	(0.46)
Parity effects					
– Parity 0	1.62				
	(0.69)				
– Parity 1	0.96				
	(0.69)				
– Parity 2	−1.02				
	(0.69)				
– Parity 3	−1.56				
	(0.69)				
N	20,018	1,339	3,816	9,503	5,360

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

References

Abío, G., Patxot, C., Sánchez-Romero, M., & Souto, G. (2017). The welfare state and demographic dividends. *Demographic Research*, *36*, 1453–1490. doi:10.4054/DemRes.2017.36.48

Abnett, K. (2021). EU launches big climate plan for "our children and grandchildren". Retrieved from https://www.reuters.com/business/environment/wrapup-eu-launches-big-climate-plan-our-children-grandchildren-2021-07-14/

Allison, P. D. (2009). *Fixed Effects Regression Models*. Thousand Oaks, CA: SAGE Publications, Inc.

Amuedo-Dorantes, C., & Kimmel, J. (2005). The motherhood wage gap for women in the United States: The importance of college and fertility delay. *Review of Economics of the Household*, *3*(1), 17–48. doi:10.1007/s11150-004-0978-9

Anderson, J. E. (1975). The relationship between change in educational attainment and fertility rates in Taiwan. *Studies in Family Planning*, *6*(3), 72–81. doi:10.2307/1964936

Anderson, T., & Kohler, H.-P. (2013). Education Fever and the East Asian Fertility Puzzle. *Asian Population Studies*, *9*(2), 196–215. doi:10.1080/17441730.2013.797293

Ang, X. L. (2015). The Effects of Cash Transfer Fertility Incentives and Parental Leave Benefits on Fertility and Labor Supply: Evidence from Two Natural Experiments. *Journal of Family and Economic Issues*, *36*(2), 263–288. doi:10.1007/s10834-014-9394-3

Angrist, J. D., & Pischke, J.-S. (2009). *Mostly harmless econometrics : an empiricist's companion.* Princeton: Princeton University Press.

Avery, R., & Freedman, R. (1970). Taiwan – Implications of Fertility at Replacement Levels. *Studies in Family Planning*, (59), 1–4.

Baizan, P., Arpino, B., & Delclos, C. E. (2016). The Effect of Gender Policies on Fertility: The Moderating Role of Education and Normative Context. *European Journal of Population*, *32*(1), 1–30. doi:10.1007/s10680-015-9356-y

Balbo, N., Billari, F. C., & Mills, M. (2013). Fertility in Advanced Societies: A Review of Research. *European Journal of Population / Revue européenne de Démographie*, *29*(1), 1–38. doi:10.1007/s10680-012-9277-y

Basten, S. (2013). Comparing Projection Assumptions of Fertility in Six Advanced Asian Economies; or 'Thinking Beyond the Medium Variant'. *Asian Population Studies*, *9*(3), 322–331. doi:10.1080/17441730.2013.836394

Basten, S., Lutz, W., & Scherbov, S. (2013). Very long range global population scenarios to 2300 and the implications of sustained low fertility. *Demographic Research*, *28*, 1145–1166. doi:10.4054/DemRes.2013.28.39

Basten, S., & Verropoulou, G. (2015). A Re-Interpretation of the 'Two-child Norm' in Post-Transitional Demographic Systems: Fertility Intentions in Taiwan. *PLOS ONE*, *10*(8). doi:10.1371/journal.pone.0135105

Baum, C. L., & Ruhm, C. J. (2016). The Effects of Paid Family Leave in California on Labor Market Outcomes. *Journal of Policy Analysis and Management*, *35*(2), 333–U135. doi:10.1002/pam.21894

Baxter, J., Hewitt, B., & Haynes, M. (2008). Life Course Transitions and Housework: Marriage, Parenthood, and Time on Housework. *Journal of Marriage and Family*, *70*(May), 259–272.

Bayne, E. (2009). Gender pedagogy in Swedish pre-schools: An Overview. *Gender Issues*, *26*(2), 130–140. doi:10.1007/s12147-009-9076-x

Beaujouan, E., & Berghammer, C. (2019). The Gap Between Lifetime Fertility Intentions and Completed Fertility in Europe and the United States: A Cohort Approach. *Population Research and Policy Review*, *38*(4), 1–29. doi:http://dx.doi.org/10.1007/s11113-019-09516-3

Becker, G. S. (1973). A Theory of Marriage: Part I. *Journal of Political Economy*, *81*(4), 813–846.

Becker, G. S. (1974). A Theory of Marriage: Part II. *Journal of Political Economy*, *82*(2), S11–S26.

Becker, G. S., Murphy, K. M., & Tamura, R. (1990). Human Capital, Fertility, and Economic Growth. *Journal of Political Economy*, *98*(5), S12–S37.

Begall, K., & Mills, M. C. (2012). The Influence of Educational Field, Occupation, and Occupational Sex Segregation on Fertility in the Netherlands. *European Sociological Review*, *29*(4), 720–742. doi:10.1093/esr/jcs051

Behrman, J. R., & Rosenzweig, M. R. (2002). Does Increasing Women's Schooling Raise the Schooling of the Next Generation? *The American Economic Review*, *92*(1), 323–334.

Bergsvik, J., Fauske, A., & Hart, R. K. (2020). *Effects of policy on fertility: A systematic review of (quasi)experiments* (Discussion Papers No. 922). Statistics Norway. Oslo.

Bloom, D. E., Canning, D., & Fink, G. (2011). *Implications of Population Aging for Economic Growth* (Working Paper No. 16705). National Bureau of Economic Research. doi:10.3386/w16705

Bloom, D. E., Canning, D., Fink, G., & Finlay, J. E. (2010). The Cost of Low Fertility in Europe. *European Journal of Population*, *26*(2), 141–158.

Boll, C., Leppin, J., & Reich, N. (2014). Paternal childcare and parental leave policies: evidence from industrialized countries. *Review of Economics of the Household*, *12*(1, SI), 129–158.

Brett, E. A. (2009). *Reconstructing Development Theory: International Inequality, Institutional Reform and Social Emancipation*. Basingstoke: Palgrave Macmillan.

Brinton, M. C., & Lee, D.-J. (2016). Gender-Role Ideology, Labor Market Institutions, and Post-industrial Fertility. *Population and Development Review*, *42*, 405–433. doi:10.1111/padr.161

Brinton, M. C., Bueno, X., Olah, L., & Hellum, M. (2018). Postindustrial Fertility Ideals, Intentions, and Gender Inequality: A Comparative Qualitative Analysis. *Population and Development Review*, *44*, 281–309. doi:10.1111/padr.12128

Burggraf, S. P. (1993). How Should the Costs of Child Rearing Be Distributed? *Challenge*, *36*(5), 48–55.

Buttner, T., & Lutz, W. (1990). Estimating Fertility Responses to Policy Measures in the German Democratic Republic. *Population and Development Review*, *16*(3), 539–555.

Cai, F. (2020). The Second Demographic Dividend as a Driver of China's Growth. *China & World Economy*, *28*(5), 26–44. doi:https://doi.org/10.1111/cwe.12350

Caldwell, J. C., Caldwell, P., & McDonald, P. (2002). Policy Responses to Low Fertility and its Consequences : A Global Survey. *Journal of Population Research*, *19*(1), 1–24.

Campbell Collaboration. (2019). Campbell Systematic Reviews: Policies and Guidelines. *Campbell Policies and Guidelines Series No. 1*. doi:10.4073/cpg.2016.1

Cannonier, C. (2014). Does the Family and Medical Leave Act (FMLA) Increase Fertility Behavior? *Journal of Labor Research*, *35*(2), 105–132. doi:10.1007/s12122-014-9181-9

Carneiro, P., Løken, K. V., & Salvanes, K. G. (2015). A Flying Start? Maternity Leave Benefits and Long-Run Outcomes of Children. *Journal of Political Economy*, *123*(2), 365–412.

Casterline, J., & Gietel-Basten, S. (2018). Exploring family demography in Asia through the lens of fertility preferences. In S. Gietel-Basten, J. Casterline, & M. K. Choe (Eds.), *Family Demography in Asia: A Comparative Analysis of Fertility Preferences* (Chap. 1, pp. 1–14). Cheltenham, UK: Edward Elgar.

Casterline, J. B., & Han, S. (2017). Unrealized fertility: Fertility desires at the end of the reproductive career. *Demographic Research*, *36*, 427–453. doi:10.4054/DemRes.2017.36.14

Cernada, G., Sun, T. H., Chang, M. C., & Tsai, J. F. (2007). Taiwan's Population and Family Planning Efforts: An Historical Perspective. *International Quarterly of Community Health Education*, *27*(2), 99–120. doi:10.2190/IQ.27.2.b

Cerrato, J., & Cifre, E. (2018). Gender inequality in household chores and work-family conflict. *Frontiers in Psychology*, *9*(AUG), 1–11. doi:10.3389/fpsyg.2018.01330

Cetre, S., Clark, A. E., & Senik, C. (2016). Happy People Have Children: Choice and Self- Selection into Parenthood. *European Journal of Population*, *32*(3), 445–473. doi:10.1007/s10680-016-9389-x

Chan, S.-J., & Lin, L.-W. (2015). Massification of Higher Education in Taiwan: Shifting Pressure from Admission to Employment. *Higher Education Policy*, *28*(1), 17–33. doi:10.1057/hep.2014.33

Chang, M. (1982). Age at marriage and fertility in Taiwan. *Jing Ji Lun Wen*, *10*(2), 129–155.

Chang, M. C. (1980). Migration and fertility in Taiwan. *Ren Kou Xue Kan*, (4), 35–69.

Chang, M.-C. (2006). Taiwan's Transition from High Fertility to Lowest Low Levels. *Asian Journal of Health and Information Sciences*, *1*(1), 1–15. doi:10.6412/AJHIS.200604.0001

Chang, M.-C., Freedman, R., & Sun, T.-H. (1981). Trends in fertility, family size preferences, and family planning practice: Taiwan, 1961-80. *Studies in Family Planning*, *12*(5), 211–228. doi:10.2307/1965830

Chen, C.-N., & Liu, P. K.-C. (2007). Is Taiwan's Lowest-low Fertility Reversible Via Socio-economic Development? *Journal of Population Studies*, (34), 1–36.

Chen, H., Chen, Y., Liao, Y., & Chen, H. (2013). Relationship of Fertility With Intelligence and Education in Taiwan: A Brief Report. *Journal of Biosocial Science*, *45*(4), 567–571. doi:10.1017/S0021932012000545

Chen, H. (2012). Trends in Low Fertility and Policy Responses in Taiwan. *The Japanese Journal of Population*, *10*(1), 78–88.

Chen, Y.-H., & Chen, H. (2014). Continuity and Changes in the Timing and Formation of First Marriage Among Postwar Birth Cohorts in Taiwan. *Journal of Family Issues*, *35*(12), 1584–1604. doi:10.1177/0192513X14538026

Chen, I.-C. (2016). Parental Education and Fertility: An Empirical Investigation Based on Evidence from Taiwan. *Journal of Family and Economic Issues*, *37*(2), 272–284. doi:10.1007/s10834-015-9448-1

Chen, M., & Yip, P. S. F. (2017). The Discrepancy Between Ideal and Actual Parity in Hong Kong: Fertility Desire, Intention, and Behavior. *Population Research and Policy Review*, *36*(4), 583–605. doi:10.1007/s11113-017-9433-5

Chen, W.-Y. (2013). Does Housing Cost Affect Birth Rates in Taiwan? The ADL Test for Threshold Co-Integration. *Romanian Journal of Economic Forecasting*, *16*(3), 90–103.

Chen, Y.-H. (2008). The significance of cross-border marriage in a low fertility society: Evidence from Taiwan. *Journal of Comparative Family Studies*, *39*(3), 331–352.

Cheng, B. S., & Nwachukwu, S. L. S. (1997). The Effect of Education on Fertility in Taiwan: A Time Series Analysis. *Economics Letters*, *56*(1), 95–99.

Cheng, K.-W. (2011). The Effect of Contraceptive Knowledge on Fertility: The Roles of Mass Media and Social Networks. *Journal of Family & Economic Issues*, *32*(2), 257–267.

Cheng, L.-H. (2004). Labor Market Structure and Fertility: a Comparison of Taiwan and South Korea. American sociological association annual meeting, San Francisco.

Cheng, S. Y., & Jacob, W. J. (2012). Expansion and stratification of higher educational opportunity in Taiwan. *Chinese Education and Society*, *45*(5-6), 112–133.

Cheng, Y. H. A. (2014). Changing partner choice and marriage propensities by education in post-industrial Taiwan, 2000-2010. *Demographic Research*, *31*(1), 1007–1042. doi:10.4054/DemRes.2014.31.33

Cheng, Y.-h. A., & Hsu, C.-H. (2020). No More Babies without Help for Whom? Education, Division of Labor, and Fertility Intentions. *Journal of Marriage and Family*, *82*(4), 1270–1285. doi:10.1111/jomf.12672

Cheng, Y.-h. A., & Loichinger, E. (2017). The Future Labor Force of an Aging Taiwan: The Importance of Education and Female Labor Supply. *Population Research and Policy Review*, *36*(3), 441–466.

Cheng, Y.-h. A., & Yang, C.-l. W. (2021). Continuity and changes in attitudes toward marriage in contemporary Taiwan. *Journal of Population Research*, *38*(2), 139–167. doi:10.1007/s12546-021-09259-z

Cheung, A. K. L., & Kim, E. H. W. (2018). Domestic Labor, Attitudes, and Women's Marital Satisfaction: A Longitudinal Study in Korea. *Journal of Family Issues*, *39*(16), 3931–3955. doi:10.1177/0192513X18800813

Chi, P., & Hsin, P. (1996). Family structure and and fertility behavior in Taiwan. *Population Research and Policy Review*, *15*(4), 327–339. doi:10.1007/BF00128428

Chow, L. P. (1965). A Program to Control Fertility in Taiwan – Setting, Accomplishment and Evaluation. *Population Studies*, *19*(2), 155–166. doi:10.2307/2173186

Chuang, H.-W., & Lin, E. S. (2006). The Evolution of the Empirical Study of Taiwan's Female Labor Supply. *Taiwan Economic Review*, *34*(2), 119–172.

Coleman, D., & Rowthorn, R. (2011). Who's Afraid of Population Decline? A Critical Examination of Its Consequences. *Population and Development Review*, *37*(s1), 217–248. doi:https://doi.org/10.1111/j.1728-4457.2011.00385.x

Collver, A., Speare, A., & Liu, P. K. C. (1967). Local Variations of Fertility in Taiwan. *Population Studies*, *20*(3), 329–342. doi:10.2307/2172676

Cooke, L. P. (2004). The Gendered Division of Labor and FamilyOutcomes in Germany. *Journal of Marriage and Family*, *66*(5), 1246–1259. doi:10.1111/j.0022-2445.2004.00090.x

Cooke, L. P. (2009). Gender Equity and Fertility in Italy and Spain. *Journal of Social Policy*, *38*(1), 123. doi:10.1017/S0047279408002584

Cools, S., Fiva, J. H., & Kirkebøen, L. J. (2015). Causal Effects of Paternity Leave on Children and Parents. *The Scandinavian Journal of Economics*, *117*(3), 801–828. doi:10.1111/sjoe.12113

Coombs, L. C. (1979). Prospective Fertility and Underlying Preferences – Longitudinal-Study in Taiwan. *Population Studies*, *33*(3), 447–455. doi:10.2307/2173891

Coombs, L. C., & Sun, T.-H. (1981). Familial Values in a Developing Society: A Decade of Change in Taiwan. *Social Forces*, *59*(4), 1229–1255.

Cygan-Rehm, K. (2016). Parental leave benefit and differential fertility responses: evidence from a German reform. *Journal of Population Economics*, *29*(1), 73–103. doi:10.1007/s00148-015-0562-z

D'Albis, H., Gobbi, P. E., & Greulich, A. (2017). Having a Second Child and Access to Childcare: Evidence From European Countries. *Journal of Demographic Economics*, *83*(2), 177–210. doi:{10.1017/dem.2017.2}

Dahl, G. B., Løken, K. V., Mogstad, M., & Vea Salvanes, K. (2013). What is the Case for Paid Maternity Leave? *Review of Economics & Statistics*, *98*(4), 655–670.

Del Bono, E., Weber, A., & Winter-Ebmer, R. (2015). Fertility and economic instability: the role of unemployment and job displacement. *Journal of Population Economics*, *28*(2), 463–478. doi:10.1007/s00148-014-0531-y

Dommermuth, L., Hohmann-Marriott, B., & Lappegård, T. (2017). Gender Equality in the Family and Childbearing. *Journal of Family Issues*, *38*(13), 1803–1824. doi:10.1177/0192513X15590686

Duvander, A.-Z., & Johansson, M. (2012). What are the effects of reforms promoting fathers' parental leave use? *Journal of European Social Policy*, *22*(3), 319–330. doi:10.1177/0958928712440201

Duvander, A.-Z., Johansson, M., & Lappegard, T. (2016). *Family Policy Reform Impact on Continued Fertility in the Nordic Countries* (Working Papers No. 2). Stockholm University Linnaeus Center on Social Policy and Family Dynamics in Europe (SPaDE).

Duvander, A.-Z., Lappegard, T., & Johansson, M. (2020). Impact of a Reform Towards Shared Parental Leave on Continued Fertility in Norway and Sweden. *Population Research and Policy Review*, *39*(6), 1205–1229. doi:10.1007/s11113-020-09574-y

Duvander, A.-Z., Lappegård, T., Andersen, S. N., Garðarsdóttir, Ó., Neyer, G., & Viklund, I. (2019). Parental leave policies and continued childbearing in Iceland, Norway, and Sweden. *Demographic Research*, *40*, 1501–1528. doi:http://dx.doi.org/10.4054/DemRes.2019.40.51

Dzenovska, D. (2011). Notes on Emptiness and the Importance of Maintaining Life. *Anthropology of East Europe Review*, *29*(2), 228–241.

Dzenovska, D. (2018). Emptiness and its futures Staying and leaving as tactics of life in Latvia. *Focaal: Journal of Global and Historical Anthropology*, (80), 16–29.

Dzenovska, D. (2020). Emptiness: Capitalism without people in the Latvian countryside. *American Ethnologist*, *47*(1), 10–26. doi:10.1111/amet.12867

Ermisch, J. F. (2003). *An Economic Analysis of the Family*. Princeton: Princeton University Press.

Esping-Andersen, G. (1990). *The Three Worlds of Welfare Capitalism*. Princeton: Princeton University Press.

Esping-Andersen, G., & Billari, F. C. (2015). Re-theorizing Family Demographics. *Population and Development Review*, *41*(1), 1–31. doi:10.1111/j.1728-4457.2015.00024.x

Esteve, A., Kashyap, R., Garcia Roman, J., Cheng, Y.-h. A., Fukuda, S., Nie, W., & Lee, H.-o. (2020). Demographic change and increasing late singlehood in East Asia, 2010-2050. *Demographic Research*, *43*, 1367–1398. doi:10.4054/DemRes.2020.43.46

Farré, L., & González, L. (2019). Does paternity leave reduce fertility? *Journal of Public Economics*, *172*, 52–66. doi:10.1016/j.jpubeco.2018.12.002

Feeney, G. (1991). Fertility Decline in Taiwan: A Study Using Parity Progression Ratios. *Demography*, *28*(3), 467–479.

Feeney, G. (2003). Momentum of Population Growth. Macmillan Reference USA.

Freedman, D. S. (1975). Consumption of Modern Goods and Services and its Relation to Fertility – Study in Taiwan. *Journal of Development Studies*, *12*(1), 95–117. doi:10.1080/00220387508421562

Freedman, R. (1965). The accelerating fertility decline in Taiwan. *Population Index*, *31*(4), 430–435.

Freedman, R. (1986). Policy Options after the Demographic Transition: The Case of Taiwan. *Population and Development Review*, *12*(1), 77–100.

Freedman, R., Coombs, L. C., Chang, M. C., & Sun, T. H. (1974). Trends in Fertility, Family Size Preferences, and Practice of Family Planning – Taiwan, 1965-1973. *Studies in Family Planning*, *5*(9), 270–288. doi:10.2307/1965133

Freedman, R., Hermalin, A., & Sun, T. H. (1972). Fertility Trends in Taiwan – 1961-1970. *Population Index*, *38*(2), 141–164. doi:10.2307/2733091

Freedman, R., Hermalin, A. I., & Chang, M. C. (1975). Do Statements About Desired Family-Size Predict Fertility – Case of Taiwan, 1967-1970. *Demography*, *12*(3), 407–416. doi:10.2307/2060824

Freedman, R., Jain, A. K., Hermalin, A. I., & Sun, T. H. (1971). Fertility After Insertion of an IUCD in Taiwans Family-Planning Program. *Social Biology*, *18*(1), 46. doi:10.1080/19485565.1971.9987899

Freedman, R., & Muller, J. (1967). Continuing Fertility Decline in Taiwan – 1965. *Population Index*, *33*(1), 3–17. doi:10.2307/2732379

Freedman, R., Peng, J. Y., Takeshita, Y., & Sun, T. H. (1963). Fertility Trends in Taiwan – Tradition and Change. *Population Studies*, *16*(3), 219–236.

Freedman, R., & Takeshita, J. Y. (1964). Changing Fertility in Taiwan. *Population Index*, *30*(3), 292.

Freedman, R., Takeshita, J. Y., & Sun, T. H. (1964). Fertility and Family-Planning in Taiwan – A Case-Study of the Demographic-Transition. *American Journal of Sociology*, *70*(1), 16–27. doi:10.1086/223734

Freedman, R., Chang, M. C., Sun, T. H., & Weinstein, M. (1994a). The Fertility Transition in Taiwan. In A. Thornton & H.-S. Lin (Eds.), *Social Change and the Family in Taiwan* (Chap. 11, pp. 264–304). Chicago: The University of Chicago Press.

Freedman, R., Chang, M.-C., & Sun, T.-H. (1994b). Taiwan's Transition from High Fertility to Below-Replacement Levels. *Studies in Family Planning*, *25*(6), 317–331.

Freedman, R., Fan, T.-H., Wei, S.-P., & Weinberger, M. B. (1977). Trends in Fertility and in the Effects of Education on Fertility in Taiwan, 1961-74. *Studies in Family Planning*, *8*(1), 11–18. doi:10.2307/1965423

Freeman, E., Xiaohong, M., Ping, Y., Wenshan, Y., & Gietel-Basten, S. (2018). 'I couldn't hold the whole thing': the role of gender, individualisation and risk in shaping fertility preferences in Taiwan. *Asian Population Studies*, *14*(1), 61–76. doi:10.1080/17441730.2017.1386408

Frejka, T., Jones, G. W., & Sardon, J. P. (2010). East Asian Childbearing Patterns and Policy Developments. *Population and Development Review*, *36*(3), 579–606. doi:10.1111/j.1728-4457.2010.00347.x

Frisco, M. L., & Williams, K. (2003). Perceived Housework Equity, Marital Happiness, and Divorce in Dual-Earner Households. *Journal of Family Issues*, *24*(1), 51–73. doi:10.1177/0192513X02238520

Gauthier, A. H. (2007). The impact of family policies on fertility in industrialized countries: a review of the literature. *Population Research and Policy Review*, *26*(3), 323–346. doi:10.1007/s11113-007-9033-x

Gauthier, A. H. (2008). Some theoretical and methodological comments on the impact of policies on fertility. *Vienna Yearbook of Population Research*, 25–28.

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models.* Cambridge: Cambridge University Press.

Gietel-Basten, S. (2018). Fertility preferences in Taiwan. In S. Gietel-Basten, J. Casterline, & M. K. Choe (Eds.), *Family Demography in Asia: A Comparative Analysis of Fertility Preferences* (Chap. 22, pp. 340–354). Cheltenham, UK: Edward Elgar.

Gietel-Basten, S. (2019). *The "Population Problem" in Pacific Asia.* Oxford: Oxford University Press.

Gietel-Basten, S., & Verropoulou, G. (2018). The changing relationship between marriage and childbearing in Hong Kong. *PLOS ONE*, *13*(3), 1–13. doi:10.1371/journal.pone.0194948

Goldscheider, F., Bernhardt, E., & Brandén, M. (2013). Domestic gender equality and childbearing in Sweden. *Demographic Research*, *29*, 1097–1126. doi:10.4054/DemRes.2013.29.40

Goldscheider, F., Bernhardt, E., & Lappegård, T. (2015). The Gender Revolution: A Framework for Understanding Changing Family and Demographic Behavior. *Population and Development Review*, *41*(2), 207–239. doi:10.1111/j.1728-4457.2015.00045.x

Goodkind, D. M. (1991). Creating New Traditions in Modern Chinese Populations: Aiming for Birth in the Year of the Dragon. *Population and Development Review*, *17*(4), 663–686. doi:10.2307/1973601

Goodkind, D. M. (1993). New Zodiacal Influences on Chinese Family Formation: Taiwan, 1976. *Demography*, *30*(2), 127–142. doi:10.2307/2061833

Greenhalgh, S. (1988). Fertility as mobility: Sinic transitions. *Population and Development Review*, *14*(4), 629–674.

Greenspan, A. (1994). After the demographic transition: policy responses to low fertility in four Asian countries. *Asia-Pacific Population & Policy*, (30), 1–4.

Grunow, D., Schulz, F., & Blossfeld, H.-P. (2012). What determines change in the division of housework over the course of marriage? *International Sociology*, *27*(3), 289–307. doi:10.1177/0268580911423056

Guo, Z., Gietel-Basten, S., & Gu, B. (2019). The lowest fertility rates in the world? Evidence from the 2015 Chinese 1% sample census. *China Population and Development Studies*, *2*(3), 245–258. doi:10.1007/s42379-018-0012-1

Gustafsson, S. (2005). Having kids later. Economic Analyses for Industrialized Countries. *Review of Economics of the Household*, *3*(1), 5–16. doi:10.1007/s11150-004-0977-x

Guttmacher. (1975). Younger women in taiwan want smaller families, fewer sons; birth control use is found widespread. *International Family Planning Digest*, *1*(4), 10–11.

Harknett, K., Billari, F. C., & Medalia, C. (2014). Do Family Support Environments Influence Fertility? Evidence from 20 European Countries. *European Journal of Population*, *30*(1), 1–33. doi:10.1007/s10680-013-9308-3

Hart, R. K., Andersen, S. N., & Drange, N. (2019). *Effects of extended paternity leave on union stability and fertility* (Discussion Papers No. 899). Statistics Norway.

Hayford, S. R. (2009). The Evolution of Fertility Expectations Over the Life Course. *Demography*, *46*(4), 765–783.

Hermalin, A. I. (1974). Empirical Research in Taiwan on Factors Underlying Differences in Fertility. *Studies in Family Planning*, *5*(10), 314–324. doi:10.2307/1965142

Hermalin, A. I., & Lavely, W. R. (1979). Agricultural-Development and Fertility Change in Taiwan. *Population Index*, *45*(3), 393.

Hermalin, A. I., & Yang, L.-S. (2004). Levels of Support from Children in Taiwan: Expectations versus Reality. *Population and Development Review*, *30*(3), 417–448.

Hermalin, A. I., Freedman, R., Sun, T.-H., & Chang, M.-C. (1979). Do Intentions Predict Fertility? The Experience in Taiwan, 1967-74. *Studies in Family Planning*, *10*(3), 75–95. doi:10.2307/1965677

HFD. (2020). *Human Fertility Database*. Max Planck Institute for Demographic Research (Germany) and Vienna Institute of Demography (Austria), 2020. Retrieved October 12, 2020, from https://www.humanfertility.org/cgi-bin/main.php

Higgins, J. P. T., & Green, S. (2008). *Cochrane Handbook for Systematic Reviews of Interventions*. Oxford: Wiley.

Hin, S., Gauthier, A., Goldstein, J., & Bühler, C. (2011). Fertility preferences: what measuring second choices teaches us. *Vienna Yearbook of Population Research*, *9*, 131–156. doi:10.1553/populationyearbook2011s131

Hiroko, T. (2005). *The Political Economy of Reproduction in Japan: Between nation-state and everyday life*. Abingdon, UK: RoutledgeCurzon.

Hoem, J. M. (1990). Social Policy and Recent Fertility Change in Sweden. *Population and Development Review*, *16*(4), 735–748.

Hoem, J. M. (1993). Public Policy as the Fuel of Fertility: Effects of a Policy Reform on the Pace of Childbearing in Sweden in the 1980s. *Acta Sociologica*, *36*(1), 19–31.

Hoem, J. M. (2008). Overview Chapter 8: The impact of public policies on European fertility. *Demographic Research*, *19*, 249–259. doi:10.4054/DemRes.2008.19.10

Hoem, J. M. (2005). Why does Sweden have such high fertility? *Demographic Research*, *13*(22), 559–572. doi:10.4054/DemRes.2005.13.22

Hong, S. H., & Sullivan, R. (2016). The Effects of Subsidies for Childbearing on Migration and Fertility: Evidence From Korea. *Singapore Economic Review*, *61*(4). doi:10.1142/s021759081550040x

Hori, M. (2017). Full-Time Employment and Marital Satisfaction among Women in East Asian Societies. *Comparative Sociology*, *16*(6), 771–787. doi:10.1163/15691330-12341444

Horton, B. P., Khan, N. S., Cahill, N., Lee, J. S. H., Shaw, T. A., Garner, A. J., … Rahmstorf, S. (2020). Estimating global mean sea-level rise and its uncertainties by 2100 and 2300 from an expert survey. *npj Climate and Atmospheric Science*, *3*(1). doi:10.1038/s41612-020-0121-5

Hsing, Y. (2003). Impact of Institutional and Socioeconomic Changes on Marital Relationship: The Case of Taiwan. *International Journal of Social Economics*, *30*(5-6), 613–618.

Hsueh, J. C. T. (2018). How Taiwan Boosted Fertility. Retrieved September 4, 2020, from https://ifstudies.org/blog/how-taiwan-boosted-fertility

Hu, C.-Y., & Kamo, Y. (2007). The Division of Household Labor in Taiwan. *Journal of Comparative Family Studies*, *38*(1), 105–124.

Huang, F.-M. (2001). Education, Earning, and Fertility in Taiwan. In *Population change and economic development in East Asia*. Stanford: Stanford University Press.

Huang, J.-T. (2003). Unemployment and Family Behavior in Taiwan. *Journal of Family and Economic Issues*, *24*(1), 27–48.

Huang, J. (2002). Personal tax exemption: The effect on fertility in Taiwan. *The Developing Economies*, *40*(1), 32–48. doi:10.1111/j.1746-1049.2002.tb00909.x

Huang, J.-T., Kao, A.-P., & Hung, W.-C. (2006). The influence of college tuition and fees on fertility rate in Taiwan. *Journal of Family and Economic Issues*, *27*(4), 626–642.

Hur, S. (2013). Expanding the Feminist Politics of Reproductionin the Context of the Developmental Hegemony of South Korea. *Asian Journal of Women's Studies*, *19*(3), 116–148. doi:10.1080/12259276.2013.11666159

Jao, Y.-H., & Li, J.-C. (2011). *Trends in the Employment of Married Mothers of Preschool-Aged Children in Taiwan*. RAND Corporation Publications Department. Santa Monica.

Jejeebhoy, S. J. (1978). Transition from Natural to Controlled Fertility in Taiwan – Cross-Sectional Analysis of Demand and Supply Factors. *Studies in Family Planning*, *9*(8), 206–211. doi:10.2307/1965865

Jejeebhoy, S. J. (1979). Supply, Demand, and Cost of Regulation Factors in the Fertility Transition in Taiwan, 1960-1973. *Population Index*, *45*(3), 418.

Jejeebhoy, S. J. (1995). *Women's Education, Autonomy and Reproductive Behaviour: Experience from Developing Countries*. Oxford: Clarendon Press.

Jia, Z. (1991). Socioeconomic development, family planning, and fertility in Taiwan. *Chinese Journal of Population Science*, *3*(2), 107–113.

Jones, G. W., & Gubhaju, B. (2009). Factors Influencing Changes in Mean Age at First Marriage and Proportions Never Marrying in the Low-Fertility

Countries of East and Southeast Asia. *Asian Population Studies*, *5*(3), 237–265.

Jones, G. W. (2007). Delayed Marriage and Very Low Fertility in Pacific Asia. *Population and Development Review*, *33*(3), 453–478. doi:https://doi.org/10.1111/j.1728-4457.2007.00180.x

Jones, G. W. (2019). Ultra-low fertility in East Asia: policy responses and challenges. *Asian Population Studies*, *15*(2), 131–149.

Kan, M.-Y., & Hertog, E. (2017). Domestic division of labour and fertility preference in China, Japan, South Korea, and Taiwan. *Demographic Research*, *36*, 557–588. doi:10.4054/DemRes.2017.36.18

Kan, M.-Y., Hertog, E., & Kolpashnikova, K. (2019). Housework share and fertility preference in four East Asian countries in 2006 and 2012. *Demographic Research*, *41*, 1021–1046. doi:10.4054/DemRes.2019.41.35

Kato, T., Kumamaru, H., & Fukuda, S. (2018). Men's participation in childcare and housework and parity progression: a Japanese population-based study. *Asian Population Studies*, *14*(3), 290–309. doi:10.1080/17441730.2018.1523977

Keng, S.-H., & Sheu, S.-J. (2011). Can Education Subsidy Reverse Falling Fertility? The Case of Taiwan. *The Journal of Development Studies*, *47*(11), 1677–1698. doi:10.1080/00220388.2011.609585

Kiger, G., & Riley, P. J. (1996). Gender Differences in Perceptions of Household Labor. *The Journal of Psychology*, *130*(4), 357–370. doi:10.1080/00223980.1996.9915024

Kim, E. H.-W. (2017). Division of domestic labour and lowest-low fertility in South Korea. *Demographic Research*, *37*, 743–768. doi:10.4054/DemRes.2017.37.24

Kim, Y. M. (2013). Dependence on Family Ties and Household Division of Labor in Korea, Japan, and Taiwan. *Asian Journal of Women's Studies*, *19*(2), 7–35. doi:10.1080/12259276.2013.11666147

Kneale, D., & Joshi, H. (2008). Postponement and childlessness - Evidence from two British cohorts. *Demographic Research*, *19*, 1935–1964. doi:10.4054/DemRes.2008.19.58

Kotsadam, A., & Finseraas, H. (2011). The state intervenes in the battle of the sexes: Causal effects of paternity leave. *Social Science Research*, *40*(6), 1611–1622. doi:10.1016/j.ssresearch.2011.06.011

Kravdal, Ø. (1992). The Emergence of a Positive Relation between Education and Third Birth Rates in Norway with Supportive Evidence from the United States. *Population Studies*, *46*(3), 459–475.

Kravdal, Ø. (2010). Demographers' interest in fertility trends and determinants in developed countries: Is it warranted? *Demographic Research*, *22*, 663–690. doi:10.4054/DemRes.2010.22.22

Kuo, M.-K., & Yue, J. C. (2008). An Empirical Study of Simulation and Stochastic Methods on Population Projections. *Journal of Population Studies*, *36*, 67–98.

Lai, M. S. (2011). When Having Many Children Pays: A Case Study from Taiwan. *Journal of Population Economics*, *25*(1), 323–348.

Lalive, R., & Zweimüller, J. (2009). How Does Parental Leave Affect Fertility and Return to Work? Evidence from Two Natural Experiments. *The Quarterly Journal of Economics*, *124*(3), 1363–1402.

Lappegård, T. (2010). Family Policies and Fertility in Norway. *European Journal of Population*, *26*(1), 99–116.

Le Moglie, M., Mencarini, L., & Rapallini, C. (2015). Is it just a matter of personality? On the role of subjective well-being in childbearing behavior. *Journal of Economic Behavior and Organization*, *117*, 453–475. doi:10.1016/j.jebo.2015.07.006

Lechner, M. (2010). The Estimation of Causal Effects by Difference-in-Difference Methods. *Foundations and Trends in Econometrics*, *4*(3), 165–224. doi:10.1561/0800000014

Lee, J., & Jeong, E. (2021). The 4B movement: envisioning a feminist future-with/in a non-reproductive future in Korea. *Journal of Gender Studies*, *30*(5), 633–344. doi:https://doi.org/10.1080/09589236.2021.1929097

Lee, M. (2009). Transition to Below Replacement Fertility and Policy Response in Taiwan. *The Japanese Journal of Population*, *7*(1), 71–86.

Lee, M., & Lin, Y.-H. (2016). Transition from Anti-natalist to Pro-natalist Policies in Taiwan. In R. R. Rindfuss & M. K. Choe (Eds.), *Low fertility, institutions, and their policies: Variations across industrialized countries* (pp. 259–282). New York: Springer.

Lee, S. H., Yen, C. H., Wu, W. Y., & Lee, M. C. (2007). A Review on Adolescent Childbearing in Taiwan: Its Characteristics, Outcomes and Risks. *Asia-Pacific Journal of Public Health*, *19*(1), 40–42. doi:10.1177/10105395070190010801

Lee, Y. J., Parish, W. L., & Willis, R. J. (1994). Sons, Daughters, and Intergenerational Support in Taiwan. *American Journal of Sociology*, *99*(4), 1010–1041. doi:10.1086/230370

Leete, R. (1994). The Continuing Flight from Marriage and Parenthood Among the Overseas Chinese in East and Southeast-Asia – Dimensions and Implications. *Population and Development Review*, *20*(4), 811–829. doi:10.2307/2137663

Lesthaeghe, R., & van de Kaa, D. J. (1986). Twee demografische transities. In R. Lesthaeghe & D. J. van de Kaa (Eds.), *Groei of krimp?* Deventer: Van Loghum Slaterus.

Lesthaeghe, R. (2020). The second demographic transition, 1986–2020: sub-replacement fertility and rising cohabitation—a global update. *Genus*, *76*(1).

Lesthaeghe, R., & Neels, K. (2002). From the First to the Second Demographic Transition: An Interpretation of the Spatial Continuity of Demographic Innovation in France, Belgium and Switzerland. *European Journal of Population / Revue européenne de Démographie*, *18*(4), 325–360. doi:10.1023/A:1021125800070

Lesthaeghe, R. (2010). The Unfolding Story of the Second Demographic Transition. *Population and Development Review*, *36*(2), 211–251. doi:10.1111/j.1728-4457.2010.00328.x

Li, T.-C., & Yang, C.-L. (2004). Transitional Patterns of Female Labor Force Participation in Taiwan. *Journal of Population Studies*, *28*, 109–134.

Li, W., Chen, S., Wang, J., Zhao, L., & Wang, D. (2019). Fertility cost in China should be borne by the state. *The Lancet Global Health*, *7*(6), E708. doi:https://doi.org/10.1016/S2214-109X(19)30155-X

Li, Z., Yang, H., Zhu, X., & Xie, L. (2020). A Multilevel Study of the Impact of Egalitarian Attitudes Toward Gender Roles on Fertility Desires in China. *Population Research and Policy Review*. doi:10.1007/s11113-020-09600-z

Lin, Y.-F., & Kamo, Y. (2015). Comparative Analysis of Determinants of Birth Rates in East Asian and Western Countries. *Population Review*, *54*(1), 27–52.

Lin, T.-c. (2009). The decline of son preference and rise of gender indifference in Taiwan since 1990. *Demographic Research*, *20*, 377–402. doi:10.4054/DemRes.2009.20.16

Lin, W.-I., & Yang, S.-Y. (2009). From Successful Family Planning to the Lowest of Low Fertility Levels: Taiwan's Dilemma. *Asian Social Work and Policy Review*, *3*(2), 95–112. doi:10.1111/j.1753-1411.2009.00027.x

Liu, J.-L., & Hsu, C.-C. (2004). Economies of Scale, Gender Discrimination, and Cost of Children. *Applied Economics Letters*, *11*(6), 377–382.

Liu, P. K. (1983). The role of education in fertility transition in Taiwan. *Jing Ji Lun Wen*, *11*(2), 219–252.

Lo, K.-T. (2012). The Crowding-out Effect of Homeownership on Fertility. *Journal of Family and Economic Issues*, *33*(1), 108–117.

Luo, M. S., Wing, E., & Chui, T. (2018). Gender Division of Household Labor in China: Cohort Analysis in Life Course Patterns. *Journal of Family Issues*, *39*(12), 3153–3176. doi:10.1177/0192513X18776457

Luppi, F. (2016). When is the Second One Coming? The Effect of Couple's Subjective Well-Being Following the Onset of Parenthood. *European Journal of Population*, *32*(3), 421–444. doi:10.1007/s10680-016-9388-y

Lupri, E., & Frideres, J. (1981). The Quality of Marriage and the Passage of Time: Marital Satisfaction over the Family Life Cycle. *The Canadian Journal of Sociology*, *6*(3), 283–305.

Ma, L. (2016). Female labour force participation and second birth rates in South Korea. *Journal of Population Research*, *33*(2), 173–195. doi:10.1007/s12546-016-9166-z

Masiano, S. P., Green, T. L., Dahman, B., & Kimmel, A. D. (2019). The effects of community-based distribution of family planning services on contraceptive use: The case of a national scale-up in Malawi. *Social Science and Medicine*, *238*, 112490. doi:10.1016/j.socscimed.2019.112490

Mason, A., & Lee, R. (2006). Reform and support systems for the elderly in developing countries: capturing the second demographic dividend. *Genus*, *62*(2), 11–35.

Matsuda, S. (2019). Young Men's Employment and Their Marriage: A Comparison among Japan, South Korea, Singapore, the UK, France, and Sweden. *Comparative Sociology*, *18*(2), 204–228. doi:10.1163/15691330-12341491

Matthews, B. J. (1999). The Gender System and Fertility: An Exploration of the Hidden Links. *Canadian Studies in Population*, *26*(1), 21–38.

Matysiak, A., & Szalma, I. (2014). Effects of Parental Leave Policies on Second Birth Risks and Women's Employment Entry. *Population*, *69*(4), 599–636.

Matysiak, A., & Vignoli, D. (2008). Fertility and Women's Employment: A Meta-analysis. *European Journal of Population / Revue européenne de Démographie*, *24*(4), 363–384. doi:10.1007/s10680-007-9146-2

McDonald, P. (2000a). Gender Equity in Theories of Fertility Transition. *Population and Development Review*, *26*(3), 427–439. doi:10.1111/j.1728-4457.2000.00427.x

McDonald, P. (2000b). Gender Equity, Social Institutions and the Future of Fertility. *Journal of Population Research*, *17*(1), 1–16.

McDonald, P. (2002). Sustaining Fertility through Public Policy: The Range of Options. *Population*, *57*(3), 417–446. doi:10.2307/3246634

McDonald, P. (2006). Low Fertility and the State: The Efficacy of Policy. *Population and Development Review*, *32*(3), 485–510. doi:10.1111/j.1728-4457.2006.00134.x

McDonald, P. (2008). Very Low Fertility: Consequences, Causes and Policy Approaches. *Japanese Journal of Population*, *6*(1), 19–23.

McDonald, P. (2009). Explanations of low fertility in East Asia: a comparative perspective. In P. Straughan, A. Chan, & G. Jones (Eds.), *Ultra-low fertility in pacific asia: Trends, causes and policy issues* (pp. 23–39). New York: Routledge.

McDonald, P. (2013). Societal foundations for explaining low fertility: Gender equity. *Demographic Research*, *28*, 981–994. doi:10.4054/DemRes.2013.28.34

McGuinness, L. A., & Higgins, J. P. (2020). Risk-of-bias VISualization (robvis): An R package and Shiny web app for visualizing risk-of-bias assessments. *Research Synthesis Methods*, *12*(1), 55–61. doi:10.1002/jrsm.1411

Meier, V., & Rainer, H. (2017). Daddy months. *Journal of Population Economics*, *30*(3), 875–892. doi:10.1007/s00148-016-0631-y

Mencarini, L., & Tanturri, M. L. (2006). High Fertility or Childlessness: Micro-level Determinants of Reproductive Behaviour in Italy. *Population*, *61*(4), 389–415.

Mencarini, L., Vignoli, D., Zeydanli, T., & Kim, J. (2018). Life satisfaction favors reproduction. The universal positive effect of life satisfaction on childbearing in contemporary low fertility countries. *PLoS ONE*, *13*(12), 1–19.

Metzler, J. J. (2017). *Taiwan's transformation : 1895 to the present.* Basingstoke: Palgrave Macmillan.

Miller, A. R. (2011). The effects of motherhood timing on career path. *Journal of Population Economics*, *24*(3), 1071–1100. doi:10.1007/s00148-009-0296-x

Mills, M., Mencarini, L., Tanturri, M. L., & Begall, K. (2008). Gender equity and fertility intentions in Italy and the Netherlands. *Demographic Research*, *18*. doi:10.4054/DemRes.2008.18.1

MOE. (2018). 2018 Education Statistical Indicators. Taipei: Ministry of Education. Retrieved from https://english.moe.gov.tw/cp-27-14504-9E20A-1.html

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ (Clinical research ed.)*, *339*, b2535. doi:10.1136/bmj.b2535

MOI. (2013). *Population Policy White Paper*. Department of Household Registration, Ministry of the Interior. Retrieved from https://www.ris.gov.tw/app/en/676

MOI. (2014). *The ROC Population Policy Guidelines*. Department of Household Registration, Ministry of the Interior. Retrieved from https://www.ris.gov.tw/app/en/676

MOI. (2020). *Population Policy Data Collection (2020)*. Department of Household Registration, Ministry of the Interior. Retrieved from https://www.ris.gov.tw/app/en/676

Monnier, A. (1990). The effects of family policies in the German Democratic Republic: a re-evaluation. *Population. English Selection*, *2*, 127–140.

Montgomery, M. R., & Casterline, J. B. (1993). The Diffusion of Fertility-Control in Taiwan – Evidence from Pooled Cross-Section Time-Series Models. *Population Studies*, *47*(3), 457–479. doi:10.1080/0032472031000147246

Mueller, E., & Cohn, R. (1977). The Relation of Income to Fertility Decisions in Taiwan. *Economic Development and Cultural Change*, *25*(2), 325–347. doi:10.1086/450951

Nagase, N., & Brinton, M. C. (2017). The gender division of labor and second births: Labor market institutions and fertility in Japan. *Demographic Research*, *36*, 339–370. doi:10.4054/DemRes.2017.36.11

Nair, N. K., & Chow, L. P. (1980). Fertility intentions and behavior: some findings from Taiwan. *Studies in Family Planning*, *11*(7-8), 255–263. doi:10.2307/1966243

Nakajima, R., & Tanaka, R. (2014). Estimating the effects of pronatal policies on residential choice and fertility. *Journal of the Japanese and International Economies*, *34*, 179–200. doi:10.1016/j.jjie.2014.07.001

Nakazawa, M. (2019). *pyramid: Draw Population Pyramid*. Retrieved from https://cran.r-project.org/package=pyramid

Narayan, P. K. (2006). Determinants of Female Fertility in Taiwan, 1966-2001: Empirical Evidence from Cointegration and Variance Decomposition Analysis. *Asian Economic Journal*, *20*(4), 393–407.

Neyer, G., & Andersson, G. (2008). Consequences of family policies on childbearing behavior: Effects or artifacts? *Population and Development Review*, *34*(4), 699–724. doi:10.1111/j.1728-4457.2008.00246.x

Ní Bhrolcháin, M., & Dyson, T. (2007). On Causation in Demography: Issues and Illustrations. *Population and Development Review*, *33*(1), 1–36.

Ochiai, E. (2011). Unsustainable Societies: The Failure of Familialism in East Asia's Compressed Modernity. *Historical Social Research*, *36*(2), 219–245.

OECD. (2019a). Family benefits public spending (indicator). Retrieved from https://doi.org/10.1787/8e8b3273-en

OECD. (2019b). Gross domestic product (gdp) (indicator). Retrieved from https://doi.org/10.1787/dc2f7aec-en

OECD. (2019c). Pf1.1: Public spending on family benefits. Retrieved from http://www.oecd.org/els/family/database.htm

OECD. (2019d). Pf2.1: Parental leave systems. Retrieved from http://www.oecd.org/els/family/database.htm

Ogden, P. E., & Hall, R. (2004). The second demographic transition, new household forms and the urban population of France during the 1990s. *Transactions of the Institute of British Geographers*, *29*(1), 88–105.

Oláh, L. S. (2003). Gendering fertility: Second births in Sweden and Hungary. *Population Research and Policy Review*, *22*(2), 171–200. doi:10.1023/A:1025089031871

Olivetti, C., & Petrongolo, B. (2017). The Economic Consequences of Family Policies: Lessons from a Century of Legislation in High-Income Countries. *Journal of Economic Perspectives*, *31*(1), 205–230. doi:10.1257/jep.31.1.205

Oppenheimer, V. K. (1994). Women's Rising Employment and the Future of the Family in Industrial Societies. *Population and Development Review*, *20*(2), 293–342.

Pailhe, A., Solaz, A., & Tanturri, M. L. (2019). The Time Cost of Raising Children in Different Fertility Contexts: Evidence from France and Italy. *European Journal of Population*, *35*(2), 223–261. doi:10.1007/s10680-018-9470-8

Parr, N. (2010). Satisfaction with life as an antecedent of fertility: Partner + Happiness = Children? *Demographic Research*, *22*, 635–662. doi:10.4054/DemRes.2010.22.21

PDR. (2019). Cairo + 25: The UN Population Commission on Population Change since the 1994 ICPD. *Population and Development Review*, *45*(2), 445–448. doi:10.1111/padr.12258

Perales, F., Baxter, J., & Tai, T.-o. (2015). Gender, justice and work: A distributive approach to perceptions of housework fairness. *Social Science Research*, *51*(May), 51–63. doi:10.1016/j.ssresearch.2014.12.010

Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M., … Duffy, S. (2006). *Guidance on the conduct of narrative synthesis in systematic reviews: A product from the ESRC methods programme*. ESRC.

Population Council. (2004). The United Nations on World Population in 2300. *Population and Development Review*, *30*(1), 181–187.

Poston, D. L. J. (2000). Social and Economic Development and the Fertility Transitions in Mainland China and Taiwan. *Population and Development Review*, *26*(Supplement), 40–60.

Poston, D. L., & Zhang, L. (2014). Taiwan's Demographic Destiny: Marriage Market and Aged Dependency Implications for the Twenty-First Century. In D. L. Poston, W. S. Yang, & D. N. Farris (Eds.), *The Family and Social Change in Chinese Societies* (pp. 265–280). doi:10.1007/978-94-007-7445-2_16

Priebe, J. (2020). Quasi-experimental evidence for the causal link between fertility and subjective well-being. *Journal of Population Economics*, *33*(3), 839–882. doi:10.1007/s00148-020-00769-3

Pronzato, C. D. (2009). Return to Work after Childbirth: Does Parental Leave Matter in Europe? *Review of Economics of the Household*, *7*(4), 341–360.

PSFD. (2019). Panel survey of family dynamics. Retrieved from https://psfd.sinica.edu.tw/web/plan_01en.htm

Qian, Y., & Sayer, L. C. (2016). Division of Labor, Gender Ideology, and Marital Satisfaction in East Asia. *Journal of Marriage and Family*, *78*(2), 383–400. doi:10.1111/jomf.12274

Raley, R. K., & Raley, R. K. (2001). Increasing fertility in cohabiting unions: evidence for the second demographic transition in the United States? *Demography*, *38*(1), 59–66.

Raute, A. (2019). Can financial incentives reduce the baby gap? Evidence from a reform in maternity leave benefits. *Journal of Public Economics*, *169*, 203–222.

Ravaneral, Z. R., Young, H., Rajulton, F., & Cho, B.-Y. (1999). Should a second demographic transition follow the first? Demographic contrasts: Canada and South Korea. *Social Indicators Research*, *47*(1), 99–118. doi:10.1023/A:1006891424461

Raybould, A., & Sear, R. (2020). Children of the (gender) revolution: A theoretical and empirical synthesis of how gendered division of labour influences fertility. *Population Studies*. doi:10.1080/00324728.2020.1851748

Raymo, J. M., Park, H., Xie, Y., & Yeung, W.-j. J. (2015). Marriage and Family in East Asia: Continuity and Change. *Annual Review of Sociology*, *41*, 8.1–8.22. doi:10.1146/annurev-soc-073014-112428

Reher, D. S. (2007). Towards long-term population decline: a discussion of relevant issues. *European Journal of Population / Revue européenne de Démographie*, *23*, 189–207. doi:10.1007/s10680-007-9120-z

Rindfuss, R. R., Choe, M. K., Tsuya, N. O., Bumpass, L. L., & Tamaki, E. (2015). Do low survey response rates bias results? Evidence from Japan. *Demographic Research*, *32*, 797–828. doi:10.4054/DemRes.2015.32.26

Rindfuss, R. R., & Choe, M. K. (2016). *Low fertility, institutions, and their policies: Variations across industrialized countries*. doi:10.1007/978-3-319-32997-0

Rindfuss, R. R., Choe, M. K., & Brauner-Otto, S. R. (2016). The Emergence of Two Distinct Fertility Regimes in Economically Advanced Countries. *Population Research and Policy Review*, *35*(3), 287–304. doi:10.1007/s11113-016-9387-z

Rindfuss, R. R., Guzzo, K. B., & Morgan, S. P. (2003). The Changing Institutional Context of Low Fertility. *Population Research and Policy Review*, *22*(5), 411–438. doi:10.1023/B:POPU.0000020877.96401.b3

Roberson, P. N. E., Lenger, K. A., Norona, J. C., & Olmstead, S. B. (2018). A Longitudinal Examination of the Directional Effects between Relationship Quality and Well-Being for a National Sample of U . S . Men and Women. *Sex Roles*, *78*, 67–80. doi:10.1007/s11199-017-0777-4

Robey, B. (1991). Economic development and fertility decline: lessons from Asia's newly industrialized countries. *Asia-Pacific Population & Policy*, (16), 1–4.

Rondinelli, C., Aassve, A., & Billari, F. C. (2010). Women's wages and childbearing decisions: Evidence from Italy. *Demographic Research*, *22*, 549–578. doi:10.4054/DemRes.2010.22.19

Rutstein, S. O. (1974). Influence of Child Mortality on Fertility in Taiwan. *Studies in Family Planning*, *5*(6), 182–188. doi:10.2307/1965368

Salles, A. (2006). The Effects of Family Policy in the Former GDR on Nuptiality and Births outside Marriage. *Population*, *61*(1/2), 141–151.

Sanchez-Romero, M. (2013). The role of demography on per capita output growth and saving rates. *Journal of Population Economics*, *26*(4), 1347–1377. doi:10.1007/s00148-012-0447-3

Scanzoni, J. (1976). Sex Role Change and Influences on Birth Intentions. *Journal of Marriage and Family*, *38*(1), 43–58. doi:10.2307/350549

Schleutker, E. (2014). Determinants of Childbearing: A Review of the Literature. *Zeitschrift für Soziologie*, *43*(3), 192–211. doi:https://doi.org/10.1515/zfsoz-2014-0303

Schoonheim, M., & Huelsken, M. (2011). Religion and fertility at the extremes: The Netherlands and Taiwan, 1950-1985. *History of the Family*, *16*(3), 267–277. doi:10.1016/j.hisfam.2011.01.001

Schunck, R., & Perales, F. (2017). Within- and between-cluster effects in generalized linear mixed models: A discussion of approaches and the xthybrid command. *The Stata Journal*, *17*(1), 89–115. doi:https://doi.org/10.1177/1536867X1701700106

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference* (3rd ed.). Boston: Houghton Mifflin Company.

Shih, Y.-H., & Wang, R.-J. (2021). Incorporating gender issues into the classroom: Study on the teaching of gender-related courses in the general curriculum of Taiwan's universities. *Policy Futures in Education*, *0*(0), 1–12. doi:10.1177/14782103211009641

Song, Y.-J., Chang, K.-S., & Sylvian, G. (2013). Why are developmental citizens reluctant to procreate? Analytical insights from Shirley Sun's Population Policy and Reproduction in Singapore and Takeda Hiroko's The Political Economy of Reproduction in Japan. *Inter-Asia Cultural Studies*, *14*(3), 481–492. doi:http://dx.doi.org/10.1080/14649373.2013.801624

Spéder, Z., & Kapitány, B. (2014). Failure to Realize Fertility Intentions : A Key Aspect of the Post-communist Fertility Transition. *Population Research and Policy Review*, *33*(3), 393–418. doi:https://doi.org/10.1007/s11113-013-9313-6

Sterne, J. A., Hernán, M. A., Reeves, B. C., Savović, J., Berkman, N. D., Viswanathan, M., … Higgins, J. P. (2016). ROBINS-I: A tool for assessing risk of bias in non-randomised studies of interventions. *BMJ (Online)*, *355*, 4–10. doi:10.1136/bmj.i4919

Stokes, C. S., & Hsieh, Y. S. (1983). Female Employment and Reproductive Behaviour in Taiwan. *Demography*, *20*(3), 313–331.

Stropnik, N., & Šircelj, M. (2008). Slovenia: Generous family policy without evidence of any fertility impact. *Demographic Research*, *19*, 1019–1058. doi:10.4054/DemRes.2008.19.26

Stuart, E. A., Huskamp, H. A., Duckworth, K., Simmons, J., Song, Z., Chernew, M. E., & Barry, C. L. (2014). Using propensity scores in difference-in-differences models to estimate the effects of a policy change. *Health Services and Outcomes Research Methodology*, *14*, 166–182. doi:10.1007/s10742-014-0123-z

Sullivan, J. M., Su, K. H., & Liu, T. H. (1976). Influence of Induced-Abortion on Taiwanese Fertility. *Studies in Family Planning*, *7*(8), 231–238. doi:10.2307/1965159

Sun, S. H.-L. (2012). *Population Policy and Reproduction in Singapore: Making Future Citizens*. Abingdon, UK: Routledge.

Sun, T. H., Lin, H. S., & Freedman, R. (1978). Trends in Fertility, Family-Size Preferences, and Family-Planning Practice – Taiwan, 1961-76. *Studies in Family Planning*, *9*(4), 54–70. doi:10.2307/1965941

Sun, T.-H. (1984). Urban-Industrial Development as a Force in Rural Fertility Change: The Case of Taiwan, Republic of China. In *Rural development and human fertility*. New York: Macmillan.

Sun, T.-H., & Ting, T. Y. (1989). Innovation-diffusion or adjustment: the case of Taiwanese fertility transition. *Ren Kou Xue Kan*, *12*, 67–89.

Suzuki, T. (2013). *Low Fertility and Population Aging in Japan and Eastern Asia*. Tokyo: Springer.

Tai, T.-o., Yi, C.-C., & Liu, C.-H. (2019). Early Marriage in Taiwan: Evidence From Panel Data. *Journal of Family Issues*, *40*(14, SI), 1989–2014. doi:10.1177/0192513X19863211

Tamm, M. (2019). Fathers' parental leave-taking, childcare involvement and labor market participation. *Labour Economics*, *59*, 184–197. doi:https://doi.org/10.1016/j.labeco.2019.04.007

Tang, Y., Liu, C., & Zhang, X. (2017). Performance associated effect variations of public reporting in promoting antibiotic prescribing practice: A cluster randomized-controlled trial in primary healthcare settings. *Primary Health Care Research and Development*, *18*(5), 482–491. doi:10.1017/S1463423617000329

Thangkasemvathana, B. (1990). *Fertility Change and Economic Development in Taiwan*. Columbia: University of Missouri.

Thomas, J., Rowe, F., & Lin, E. S. (2019). The effect of parental leave policies on increasing fertility: A systematic review. National Institute for Health Research. PROSPERO International prospective register of systematic reviews. Retrieved from https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42019128493

Thomson, E., McDonald, E., & Bumpass, L. L. (1990). Fertility Desires and Fertility: Hers , His , and Theirs. *Demography*, *27*(4), 579–588.

Torr, B. M., & Short, S. E. (2004). Second Births and the Second Shift: A Research Note on Gender Equity and Fertility. *Population and Development Review*, *30*(1), 109–130. doi:10.1111/j.1728-4457.2004.00005.x

Toulemon, L. (2011). Should governments in Europe be more aggressive in pushing for gender equality to raise fertility? The first "YES". *Demographic Research*, *24*, 179–200. doi:10.4054/DemRes.2011.24.7

Trewinnard, K. (1998). Family planning in Taiwan. *The British Journal Of Family Planning*, *24*(2), 78–80.

Tsai, P.-Y. (2012). The Transformation of Leave Policies for Work-Family Balance in Taiwan. *Asian Women*, *28*(2), 27–54.

Tu, E. J.-C., Yan, Y., & Zhao, J. (2017). Ultra-low fertility, gender equity and policy considerations. *Asian Education and Development Studies*, *6*(2), 112–124. doi:10.1108/AEDS-02-2016-0016

Ugur, Z. B. (2020). Does Having Children Bring Life Satisfaction in Europe? *Journal of Happiness Studies*, *21*(4), 1385–1406. doi:10.1007/s10902-019-00135-5

UN. (2019). *World population prospects 2019: Highlights*. Department of Economic and Social Affairs, United Nations.

UNPD. (2020). *wpp2019: World Population Prospects 2019*. Retrieved from https://cran.r-project.org/package=wpp2019

Van Bavel, J. (2010). Choice of Study Discipline and the Postponement of Motherhood in Europe: The Impact of Expected Earnings, Gender Composition, and Family Attitudes. *Demography*, *47*(2), 439–458.

VanLaningham, J., Johnson, D. R., & Amato, P. (2001). Marital Happiness, Marital Duration, and the U-Shaped Curve: Evidence from a Five-Wave Panel Study. *Social Forces*, *78*(4), 1313–1341.

Wade, R. (2004). *Governing the market: economic theory and the role of government in East Asian industrialization.* Princeton: Princeton University Press.

Wang, Y.-C. (1985). *Economic Analysis of Fertility Behavior in Taiwan* (State University of New York - Stony Brook).

Wickberg, E. (1975). The Taiwan Peasant Movement, 1923-1932: Chinese Rural Radicalism Under Japanese Development Programs. *Pacific Affairs, University of British Columbia*, *48*(4), 558–582.

Wing, C., Simon, K., & Bello-Gomez, R. A. (2018). Designing Difference in Difference Studies: Best Practices for Public Health Policy Research. *Annual Review of Public Health*, *39*, 453–469. doi:10.1146/annurev-publhealth-040617-013507

Wood, J., Kluesener, S., Neels, K., & Myrskyla, M. (2020). Shifting links in the relationship between education and fertility. *Population, Space and Place*, *26*(8). doi:10.1002/psp.2342

Wood, J., & Neels, K. (2019). Local Childcare Availability and Dual-Earner Fertility: Variation in Childcare Coverage and Birth Hazards Over Place and Time. *European Journal of Population*, *35*, 913–937. doi:10.1007/s10680-018-9510-4

Wood, J., Neels, K., & Vergauwen, J. (2016). Economic and Institutional Context and Second Births in Seven European Countries. *Population Research and Policy Review*, *35*(3), 305–325. doi:10.1007/s11113-016-9389-x

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data* (2nd ed.). Cambridge, MA: MIT Press.

Wu, H. Y. (1970). A demographic study on the relationships of nuptiality, child mortality, and attitude toward fertility to actual fertility in Hsueh-Chia township in Taiwan. *Journal Of The Formosan Medical Association*, *69*(5), 243–255.

Yamanaka, K., Chang, H. C., & Lorenz, F. O. (1982). Modernity and Fertility Preference in Taiwan. *Sociological Quarterly*, *23*(4), 539–551. doi:10.1111/j.1533-8525.1982.tb01028.x

Yang, H. (2000). Education, Married Women's Participation Rate, Fertility and Economic Growth. *Journal of Economic Development*, *25*(2), 101–118.

Yang, L. S. (1981). Labor force participation and fertility of married women in Taiwan area. *Journal of Population Studies*, *5*, 119–145.

Yang, W.-Y. (2016). Differences in Gender-Role Attitudes between China and Taiwan. *Asian Women*, *32*(4), 73–95.

Yang, W. S. (2019). Evaluating the Impact of Taiwan's Fertility Policy. Retrieved September 3, 2020, from https://taiwaninsight.org/2019/03/29/evaluating-the-impact-of-taiwans-fertility-policy/

Yen, C. H., & Wang, C. M. (1973). Studies in family planning. 7. Taiwan. *Studies in Family Planning*, *4*(5), 118–123.

Yen, E. C., Yen, G., & Liu, B. C. (1989). Cultural and Family Effects on Fertility Decisions in Taiwan, ROC – Traditional Values and Family-Structure are as Relevant as Income Measures. *American Journal of Economics and Sociology*, *48*(4), 415–426. doi:10.1111/j.1536-7150.1989.tb02128.x

Yip, P. S. F., & Chen, M. (2016). An elasticity analysis of the effectiveness of pronatalist measures in Taiwan. *Asian Population Studies*, *12*(3), 273–293. doi:10.1080/17441730.2016.1221207

Yoon, S.-Y. (2016). Is gender inequality a barrier to realizing fertility intentions? Fertility aspirations and realizations in South Korea. *Asian Population Studies*, *12*(2), 203–219. doi:10.1080/17441730.2016.1163873

Yoon, S.-Y. (2017). The influence of a supportive environment for families on women's fertility intentions and behavior in South Korea. *Demographic Research*, *36*, 227–254. doi:10.4054/DemRes.2017.36.7

Yu, J., & Xie, Y. (2011). The Varying Display of "Gender Display". *Chinese Sociological Review*, *44*(2), 5–30. doi:10.2753/CSA2162-0555440201

Yu, R.-r., & Liu, Y.-s. (2014). Change and Continuity in the Experience of Marriage in Taiwan. In *Wives, Husbands, and Lovers: Marriage and Sexuality in Hong Kong, Taiwan, and Urban China*. Stanford, CA: Stanford University Press.

Zhang, L. (2011). Male and Female Fertility in Taiwan: Trends and Transitions, 1975-2004. In *Male Fertility Patterns and Determinants* (Vol. 27, pp. 57–79). doi:10.1007/978-90-481-8939-7_5